Using Visual Change Detection to Examine the Functional Architecture of Visual Short-Term Memory

Alex Burmester

A thesis submitted for the degree of Doctor of Philosophy at The University of Queensland in October 2009

Department of Human Movements Studies

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis. I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award. I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the General Award Rules of The University of Queensland, immediately made available for research and study in accordance with the Copyright Act 1968. I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material.

Statement of Contributions to Jointly Authored Works Contained in the Thesis

No jointly-authored works

Statement of Contributions by Others to the Thesis as a Whole

Guy Wallis contributed to the design of some of the experiments. Guy Wallis also suggested corrections and textual modifications to the thesis.

Statement of Parts of the Thesis Submitted to Qualify for the Award

of Another Degree

None

Published Works by the Author Incorporated into the Thesis

None

Additional Published Works by the Author Relevant to the Thesis but not Forming Part of it

None

Acknowledgements

I would like to thank my supervisor Dr. Guy Wallis for his endless patience, support and encouragement. His positive attitude has been unfailing and his willingness to impart knowledge and to offer advice has been constant and consistent. I also need to thank him for his help in my obtaining financial support for the project, without which I wouldn't have been able to start or complete it. I would like to thank my parents for their constant support, encouragement and patience. I would like to thank some special friends for helping make the ride easier - Martin, Matt, Julia, Anky, Andrew, Kim, Emilia, Lynette, Ellen, Juana, Dave, Ben Broman, Olav, Lene, Kristin, Ginger, Maria, Sheavaun, Rosey, Sam, Tabs, Abby and the two Freds. I would like to thank *Tool*, *Nine Inch Nails*, *Static-X*, *Alice in Chains*, *Opeth* and various other bands for providing a soundtrack to much of the PhD. I would like to thank Stuart and Andrew for giving me a job when my scholarship ran out and picking me up when I was at a loose end in terms of what I wanted to do next. I would like to thank Microsoft for creating the XBOX, Bungee for creating the *Halo* games and Bioware for creating *Knights of the Old Republic* and *Jade Empire*. I would like to thank Richard Garriot for creating Origin, and Origin for creating *Ultima VII* Parts I and II (the best games ever made!), *Ultima VIII* and *Ultima IX*. I would like to thank Arthur Guinness for creating Guinness. Finally, I would like to thank the other members of the lab for the fun, relaxing and positive working environment they have created.

Abstract

A common problem in vision research is explaining how humans perceive a coherent, detailed and stable world despite the fact that the eyes make constant, jumpy movements and the fact that only a small part of the visual field can be resolved in detail at any one time. This is essentially a problem of integration over time - how successive views of the visual world can be used to create the impression of a continuous and stable environment. A common way of studying this problem is to use complete visual scenes as stimuli and present a changed scene after a disruption such as an eye movement or a blank screen. It is found in these studies that observers have great difficulty detecting changes made during a disruption, even though these changes are immediately and easily detectable when the disruption is removed. These results have highlighted the importance of motion cues in tracking changes to the environment, but also reveal the limited nature of the internal representation. Change blindness studies are interesting as demonstrations but can be difficult to interpret as they are usually applied to complex, naturalistic scenes. More traditional studies of scene analysis, such as visual search, are more abstract in their formulation, but offer more controlled stimulus conditions. In a typical visual search task, observers are presented with an array of objects against a uniform background and are required to report on the presence or absence of a target object that is differentiable from the other objects in some way. More recently, scene analysis has been investigated by combining change blindness and visual search in the 'visual search for change' paradigm, in which observers must search for a target object defined by a change

over two presentations of the set of objects. The experiments of this thesis investigate change blindness using the visual search for change paradigm, but also use principles of design from psychophysical experiments, dealing with detection and discrimination of basic visual qualities such as colour, speed, size, orientation and spatial frequency. This allows the experiments to precisely examine the role of these different features in the change blindness process. More specifically, the experiments are designed to look at the capacity of visual short-term memory for different visual features, by examining the retention of this information across the temporal gaps in the change blindness experiments. The nature and fidelity of representations in visual short-term memory is also investigated by manipulating (i) the manner in which featural information is distributed across space and objects, (ii) the time for which the information is available, (iii) the manner in which observers must respond to that information. Results point to a model in which humans analyse objects in a scene at the level of features/attributes rather than at a pictorial/object level. Results also point to the fact that the working representations which humans retain during visual exploration are similarly feature- rather than object-based. In conclusion the thesis proposes a model of scene analysis in which attention and vSTM capacity limits are used to explain the results from a more information theoretic standpoint.

Keywords

vision, perception, attention, vSTM, change, blindness, detection, visual, search

Australian and New Zealand Standard Research Classifications (ANZSRC)

170112 Sensory Processes, Perception and Performance 80%, 170201 Computer Perception, Memory and Attention 15%, 170101 Biological Psychology (Neuropsychology, Psychopharmacology, Physiological Psychology) 5%

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The process of seeing begins with an observer receiving information conveyed by the light surrounding them. The study of visual perception looks at how this information is received and how it is used by an observer.

1.1 Information Processing Approaches to Visual Perception

Information Processing approaches to visual perception typically attempt to describe what is being computed from the visual input (light) and for what purpose it is being computed. They usually describe a series of stages which each involve a transformation of the input (from a previous stage, or from the eye) into a new representation.

A popular information processing approach was developed by David Marr (see Marr, 1982). It suggests the earliest stage of vision extracts the basic features of the visual scene - where boundaries, edges and borders are located (e.g., see Figure 1.1) and where regions of uniform intensity are. Further, it suggests that features such as the perspective of depth and the motion of surfaces are extracted in a later information processing stage. It suggests that three dimensional representations of objects are extracted in the final stage.

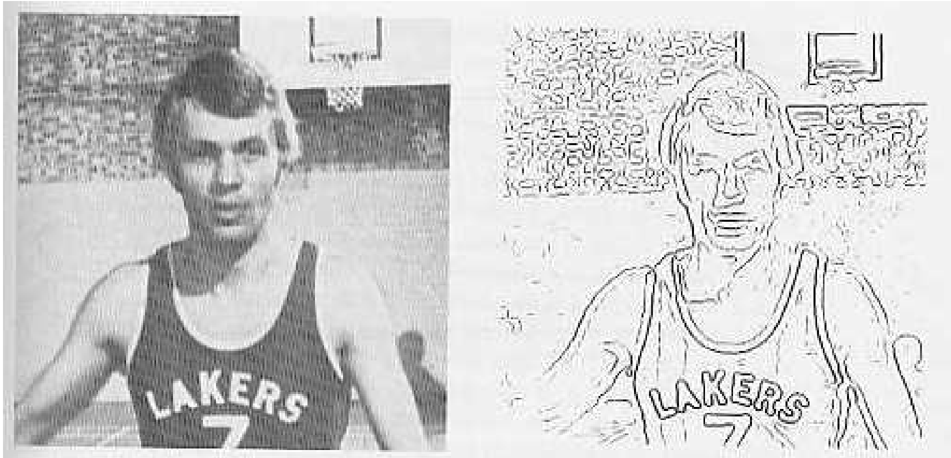


Figure 1.1: A photograph and the output generated by an algorithm used in Marr's theory of vision. The algorithm uses the second derivative of intensity variation in the image to generate boundaries corresponding to contrast changes.

1.2 Information Processing and Scene Perception

The general idea of the visual system extracting the coarse features of a scene before extracting the detailed features is common in the literature regarding the perception of visual scenes. Typically, it is suggested that the initial stage is automatic and the later stage requires conscious effort. A way to test such an approach is to present a computer-controlled scene to observers and manipulate parts of it that would be processed by the first stage and parts that would be processed by the second stage, and see if there is a difference in performance between the two conditions. This was done by Aginsky and Tarr (2000) using the *change blindness* paradigm, in which a two-dimensional photographic scene is presented followed by a blank screen and then a changed version of the scene is presented.

Aginsky and Tarr (2000) changed either the colour of an object in the scene, its presence/absence or its position (see Figure 1.2 for an example of a presence/absence change in their experiment). On average, observers were faster at detecting colour changes than the other changes, but only when the changes were cued by a word presented at the start of the trial identifying the change type (i.e., ‘colour’,‘presence/absence’ or ‘position’). The authors interpreted this as meaning that configural changes are so salient that they are not facilitated by additional information, but colour changes are not as salient and so are facilitated by this information. They proposed the reason behind this was that configural properties are extracted in a quick, automatic first-stage of scene perception whereas surface properties (such as colour) are extracted in the second stage which takes longer and requires conscious effort.

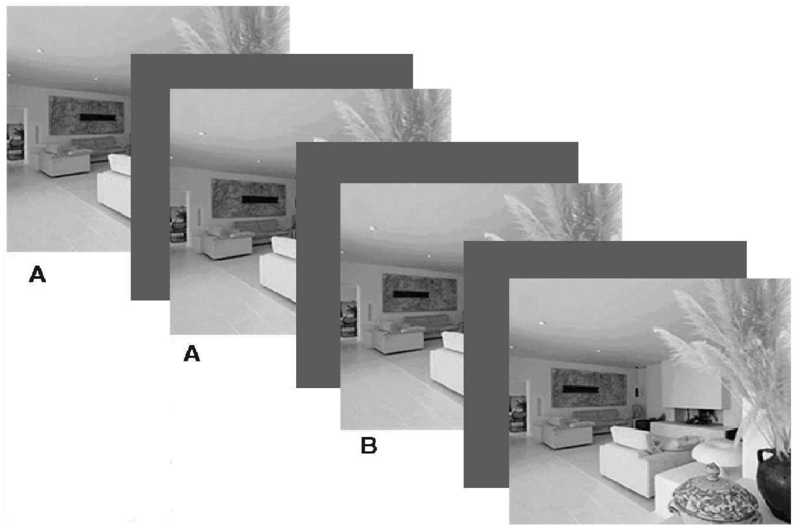


Figure 1.2: The change blindness paradigm used in the Aginsky and Tarr (2000) study. In this example, the black bar on the wall changes position from scene A to scene B.

1.3 The Thesis

The change blindness paradigm provides a way to use realistic visual stimuli and obtain information about how quickly different visual features are processed. However, modifying the change that takes place in a complex scene cannot easily be done in any quantifiable and precise manner. The visual search paradigm, in which observers search an array of objects on a computer screen for one or more targets, provides a more controllable experimental situation. The targets in the visual search paradigm are typically different from the other objects on one or more feature dimensions such as colour, orientation or shape. The visual search and change blindness paradigms can be combined by the search target being defined by a change over time rather than a difference in its visual properties to all the other elements at any point in time (see Rensink, 2000d for an example). Combining these paradigms allows us to better delineate the processes involved in change blindness.

The series of experiments in this thesis combines the visual search and change blindness paradigms to examine the way in which different features are processed by the visual system. By using the change blindness paradigm, the thesis aims to look at a higher level of processing - one that involves conscious effort, selective attention and memory, rather than the largely automatic processing that takes place in cortical areas such as V1 and V2. This is done so that the conclusions drawn from the empirical work in this thesis can be applied to the perception and analysis of complete scenes. To this end, the thesis uses the perception of scenes as a starting point in the next chapter and also as an ongoing context in which to ground the topics of different chapters.

Chapter 2

The Perception and Analysis of Visual Scenes

2.1 What is a 'Visual Scene?'

In the previous chapter, the *change blindness* paradigm was introduced. This paradigm typically uses photographs of real-world scenes and so is part of the larger field of *scene perception*. To examine the literature on scene perception, it is first necessary to provide a definition of what constitutes a visual scene. A visual scene can be defined as:

> "A semantically coherent (and often nameable) human-scaled view of a real-world environment comprising background elements and multiple discrete objects arranged in a spatially licensed manner" (Henderson, 2005) p. 849

This definition places a number of constraints on what can be classified as a visual scene. The semantic coherence constraint refers to the objects and background of the scene belonging together in a meaningful way (i.e., it is believable that this scene could occur in the real world). The human-scaling constraint means that the scene is taken from a human point of view and the size (in the visual field) of elements in the scene are as they would be from this point of view. The spatial licensing constraint refers to adherence to physical laws of gravity[1], space

[1]Although it is possible to have scenes of zero-gravity situations, typically these would not be used. If/when they are used, however, the spatial licensing constraint would presumably be

and time as well as semantic constraints imposed by object identity and function. Biederman, Mezzanotte, and Rabinowitz (1982) give the following examples of the latter constraint: a fire hydrant does not belong on top of a mailbox and a couch is not typically found outdoors. Generally, semantic constraints mean that the parts of a scene appear to belong together to a typical human observer. Although the definition of a scene given by Henderson (2005) is quite broad, it provides a starting point for a discussion of how scenes are perceived and analysed and how meaning can be attributed to them.

In scene perception research, scenes are typically presented to observers as two-dimensional photographs on a computer screen. This is done in preference to having observers view parts of the natural world because it offers better experimental control and allows scope for the study of interactions between the way the observer is viewing the scene and their performance on, say, a change detection task (e.g., Wallis & Bülthoff, 2001). Experimental control can also be enhanced by creating artificial visual scenes via computer (see Figure 2.1 for an example) - creating such scenes can allow the experimenter to control high-level features of the scene, such as the objects it contains and the environment it depicts. Also, low level properties of natural scenes can be controlled to some degree by applying statistical algorithms to the pictures that equalize them along dimensions such as spatial frequency[2] (see Johnson and Olshausen (2003) for an example of this). This chapter of the thesis deals with overt behaviours associated with scene explorations, and looks at aspects of the scene that are thought to control or modulate this behaviour.

---

relaxed from the point of view of the observer.

[2]To equalise images in the spatial frequency domain, Fourier Analysis is used. This involves extracting component sinusoidal functions from any complex periodic function. A two-dimensional image can be considered a complex periodic function, and so has component sinusoidal functions (visualised as gratings of various frequencies and orientations).

Figure 2.1: An example of a computer generated scene. Taken from http://eyelab.msu.edu/visualcognition/3d/

2.2 Viewing Visual Scenes

On the retina, photoreceptors are most densely packed in the *fovea*, an area that receives light from 2° of the visual field. Acuity is best for the fovea and falls off rapidly around it as photoreceptor density decreases (Riggs, 1965). Therefore, to get information from a visual scene, observers must move their eyes to different parts of it, so that the light reflected from those parts of the scene can be transmitted to the fovea. The voluntary eye-movements involved in scene exploration are called *saccades*. Most saccades are ballistic movements, meaning they are not modified once executed, and their duration is proportional to their amplitude, measured in degrees of visual angle (Carpenter, 1988). There are other types of movements the eyes can make (e.g., smooth pursuit, vergence, slow drifts, microsaccades, stabilisation reflexes), but it is unlikely that such movements reflect the operation of higher-level processes which are most relevant for scene perception (Henderson, 1999). Therefore, studies looking at eye-movements in scene perception have looked exclusively at saccadic eye movements.

2.2.1 Saccades and the Retinal Image

During a saccadic eye movement, the retinal image will blur. However, this blur will not be perceived due to a neurological mechanism acting to suppress retinal input. This is known as *saccadic suppression.* After each eye movement, the retinal image changes significantly, but the perceived positions of objects in the environment remain constant. In other words, objects are always perceived as being in a constant direction relative to the observer. This phenomenon is known as *visual direction constancy.* The relative position of objects could theoretically be calculated using the position of the eyes in the head, the position of the head relative to the body and the position of the object's image on the retina. Clearly, the position of the object's image on the retina comes from retinal receptors. For the source of eye position information, however, there have been two proposals. Sherrington (1906) proposed the *inflow theory*, which suggests afferent signals coming from the eye muscles are used, while Helmholtz (1866) proposed the *outflow* theory, where the command signal from the brain to the eye muscles is used instead. Evidence favours the outflow theory (see Stevens et al. (1976), Holst and Mittelstaedt (1950) and Sperry (1950)).

The preceeding discussion has highlighted two major points: that visual input is limited during saccades and that the position of the eyes is important in determining the relative position of objects in the visual field. Therefore, the placement of fixations and the time each one lasts are the major considerations of studies looking at eye movements in scene perception.

2.2.2 Fixation Placement

In the mid-1900s Alfred Yarbus ran a series of experiments in which participants viewed photographs freely or viewed them in order to answer questions posed by the experimenter (Yarbus, 1967). During these tasks, the participants' eye movements were recorded using an apparatus in which a source of light was directed towards each eye and was reflected onto a film by a small mirror attached to the surface of each eye. The apparatus was arranged such that the film recorded the position of the eye as it moved over the stimulus (see Figure 2.2). From his studies, Yarbus made this qualitative conclusion (fixation data were not analysed in any statistical way):

> "The human eyes voluntarily and involuntarily fixate on those elements of an object which carry or may carry essential and useful information. The more information is contained in an element the longer the eyes stay on it. The distribution of points of fixation on the object changes depending on the purpose of the observer." (Yarbus, 1967) p. 211

It is clear from this quote that Yarbus concluded that both the information in the scene and the task of the observer were important in determining which parts of the scene they selected as targets for fixation and which parts they fixated on for the longest period.

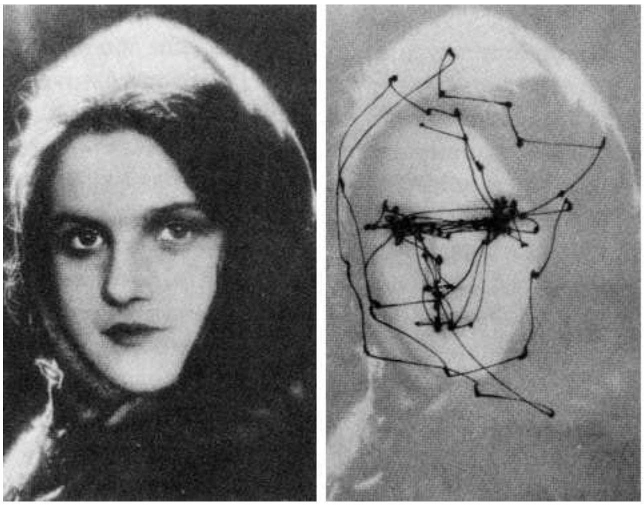


Figure 2.2: An example of a stimulus used by Yarbus (1967) and the corresponding eye movement recordings.

Mackworth and Morandi (1967) explored the relationship between the informativeness of regions in a scene and the frequency of fixations they received in a more quantitative and analytical way. Each of two colour photographs were divided into 64 regions and a group of viewers rated each region based on how easy it would be to recognise at a later viewing. These ratings were used to rate each part of each photograph on a dimension called 'informativeness', with regions judged as more recognisable being more informative. A second group had their eye movements recorded while they viewed both photographs with instructions to decide which one they preferred. The number of fixations each region received was found to be related to its informativeness. The pictures used in the Mackworth and Morandi (1967) experiment were relatively simple, but Antes (1974) replicated the effect using more complex stimuli. These studies show that informativeness of regions, as rated by one group of observers, correlates with the density of eye fixations made by another group. However, it is unclear whether the informativeness of these regions is based on their visual appearance or their semantic meaning, as informativeness was based only on the subjective judgements of a group of observers.

A study by Graef, Christiaens, and d'Ydewalle (1990) failed to replicate a finding by Loftus and Mackworth (1978) that informative objects were fixated before uninformative ones, instead finding that observers were equally as likely to fixate the two different object types (semantically informative and uninformative) in the first eight fixations of a scene, while after eight fixations they are more likely to fixate the uninformative objects. The results of Graef et al. (1990) were backed up by Henderson (1999), who also found that observers were equally likely to fixate informative and uninformative targets after the first (or second) saccade in a scene. From these studies, it appears unlikely that semantic features control initial fixations in a scene.

Mannan, Ruddock, and Wooding (1995) provided more direct evidence that *perceptual* features alone (as opposed to semantic features or semantic and perceptual features) control initial fixations - fixation positions were found to be similar for low-pass and unfiltered scenes, especially within the first 1.5 seconds of viewing. Considering that observers were unable to describe the semantic content of the low-pass scenes, this suggests perceptual features alone were controlling the placement of fixations.

2.2.3 Fixation Times

Henderson (2003, 1999) gives the average fixation duration during scene viewing as 330 msec, but states there is much variability around this point. Fixation durations form a skewed distribution varying from less than 50 msec to more than 1000 msec, with a mode around 230 msec (Henderson & Hollingworth, 1998). Several studies have looked at the question of what affects fixation duration. For instance, Loftus and Mackworth (1978) and Friedman (1979) found *first pass gaze duration* (the sum of fixation times from first entry to the exit of a particular region) was longer for semantically informative objects, while Henderson (1999) found both first and second pass gaze durations were longer for semantically informative than uninformative objects. These results show a clear effect of semantic informativeness of a region on fixation duration in that region.

2.3 Eye Movements and Attention

The studies of eye movements and scene perception discussed above look at overt behaviour in the context of scene viewing. They suggest that fixation behaviour is governed by visual and semantic information. The question is, how is this information processed and used by the observer so that it can guide their overt (motor) behaviour? In the next chapter, this issue is discussed by focussing on *visual attention*, a process that controls how visual information is selected and used by an observer to subserve their ongoing behaviour.

It is necessary to consider attention separately from eye movements, because although eye movements can give a general indicator of where an observer is looking, it has been known since Helmholtz (1866) that observers can attend to information in a visual scene without looking at it (i.e., covertly 'out of the corner of the eye'). Indeed, several researchers have suggested that eye-movements are subservient to visual attention in that they are guided by its focus and follow its movement. It is only with conscious effort (or deliberate paralysis of the eyes - see Stevens et al. (1976)) that attention can be directed without accompanying movement of the eyes.

Chapter 3

Visual Attention

3.1 What is Attention?

Typically, a visual scene in our natural environment contains a large amount and a wide variety of information at any one time. However, only a small portion of this information will be relevant to the observer's task at hand - it is the selection of this task-relevant information against a background of task-irrelevant information in a visual scene that defines the process of visual attention (Desimone & Duncan, 1995).

The general populace has a concept of attention suggesting that it is some force or quantity that can be devoted ('paid') to various stimuli or withheld from them (i.e., 'ignoring'). This 'folk psychology' treatment of attention suggests that it relies on the volition or will of the observer, but it is important to note that attention can be attracted involuntarily as well by, for example, loud noises in the auditory field or fast movements in the visual field (see H. E. Pashler (1998) for a more thorough discussion of the folk psychology of attention). Attention was originally defined in an academic context by William James as follows:

> "It is the taking possession in the mind, in clear and vivid form, of one out of several simultaneous possible objects or trains of thought." (James, 1890/1950).

This definition is useful even today, because it identifies the phenomenon of *selec-*

tivity - something that is universally accepted as being integral in defining what is known by the term 'attention' (Hatfield, 1998).

The first information-processing accounts of attention were developed in the 1950s. The best known of these was developed by Donald Broadbent (Broadbent, 1958), which characterised attention as a *selective filter* that selects some sensory inputs and rejects others (see Figure 3.1).

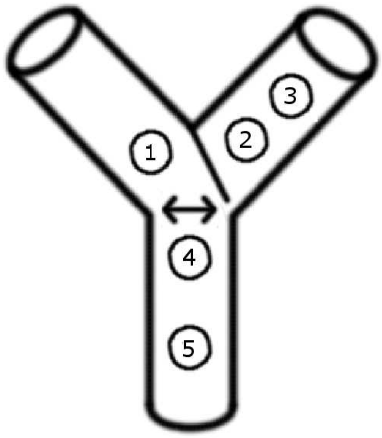


Figure 3.1: A filter-like depiction of Broadbent's early theory of attention. The numbered circles represent independent, incoming stimuli. These stimuli come through sensory channels (the holes and branches at the top) and are selected or rejected at the bottleneck, where those that are selected enter the single channel. Adapted from Broadbent (1958).

3.2 Selection and Capacity

3.2.1 Perception and Identification

Broadbent's theory makes a distinction between *perception* and *identification*, where perception involves the explicit representation of a stimulus' physical properties (e.g., its colour, luminance, spatial frequency) while identification involves matching the stimulus with a stored representation in memory and/or attaching some meaning to it. The main proposals of the theory are:

- All incoming stimuli are processed by perceptual processes, but only those that are attended to are processed to the extent where they are identified

- Attention acts as a selective filter, allowing only a subset of perceptual information to be identified

- The mechanism that identifies stimuli is capable of processing only one stimulus at a time

3.2.2 Early vs. Late Selection

Broadbent's theory is known as an *early selection* theory because it proposes that attention acts *early* in the processing stream to select stimuli for further processing. The theory is sometimes referred to as *single channel theory* because sensory stimuli are processed in multiple channels until they reach the locus of attention (the selective filter), where they must enter a *single channel*. At this point, some stimuli are selected by the attentional filter while others are prevented from entering the attentional processing stage. This filter would be located *between* the stages of perception and identification. For *late selection* theories, by contrast, the filter is located *after* both perceptual and identification processes have occurred. Therefore, identification and recognition of stimuli occur automatically and in parallel (multiple items being processed simultaneously). It is important to note, however, that most late selection theories do not suggest all available stimuli are processed beyond the perceptual stage. Instead, they suggest that all *highly familiar* stimuli are (see Deutsch and Deutsch (1963) and MacKay (1973) for examples of late selection theories).

3.2.3 Controlled Parallel Processing

Early and late selection theories are often taken as being two ends of a continuum of attention theories, varying on where they propose selection takes place in the processing stream. However, H. E. Pashler (1998) suggests a different dimensionality for categorising attention theories, defined by two questions:

- Is parallel processing (for identification) possible? or How are multiple attended stimuli processed, serially or in parallel?
- Is perceptual gating possible? or Are unattended stimuli identified?

Early selection suggests attended stimuli are processed serially, while late selection suggests they are processed in parallel. Early selection suggests unattended stimuli are not identified, while late selection suggests they are. This allows space for a theory suggesting attended stimuli are processed in parallel but unattended stimuli are not identified. H. E. Pashler (1998) terms this *controlled parallel processing* (CPP), as it selects the stimuli to be processed (in parallel) for identification. In other words, it controls the parallel processing of stimuli, as it selects which stimuli are selected for processing - and all processing occurs in parallel. There are several other theories that, like CPP, are based on a compromise between early and late selection.

3.2.4 Attenuation Theory

Introducing such a compromise, Treisman (1960) suggested that instead of being completely blocked, unattended stimuli could be attenuated. One of the central ideas of this *attenuation theory* was that recognition takes place through accumulation of information in detection units (i.e., early in the processing stream). Therefore, a stimulus is only identified if sufficient perceptual information (i.e., a threshold) has accumulated for it. Not attending to a stimulus will reduce the amount of information accumulated.

3.2.5 Capacity Sharing

Like attenuation theory, *graded capacity sharing* theory compromises between early- and late- selection. This theory proposes that resources available to process stimuli are finite and that this processing capacity is shared amongst perceptual processes (i.e., stimuli). Therefore, recognition of a stimulus could take longer/be less accurate if the observer is attending to multiple stimuli. In this scheme, processing will proceed in parallel as long as there are enough resources available, but will become serial when resources are depleted. Furthermore, the locus of attention (i.e., the point in the processing stream where selection takes place) would vary depending on the number of stimuli being attended and the resources available. The presence of limited processing resources results in *attentional capacity limits*, which are present when the speed or quality of processing decreases when additional stimuli are attended to. In other words, if the processing of a stimulus is independent of the processing of other stimuli, no capacity limits are present.

3.3 Visual Attention in Search Tasks

3.3.1 The Search Task

The study of visual attention gained momentum through the development of the *visual search* paradigm. In a typical visual search experiment, an observer is presented with an array of small stimuli such as letters, numbers or small shapes. This array will contain one or more 'target' stimuli amongst 'distractor' stimuli. On each trial, an observer is asked to indicate whether or not a target stimulus is present, typically by pressing one of two buttons. The dependent measure can either be the accuracy of the observer's judgement (how many times they correctly say the target is or is not present when the search field is presented for a fixed time on each trial), or their reaction time (RT), in which case the search field is present until the observer responds 'present' or 'absent', usually through a button press. Data from a visual search experiment is conventionally plotted as set size against accuracy or reaction time (see Figure 3.4). The slope of these functions are typically taken to show how efficient the search process is.

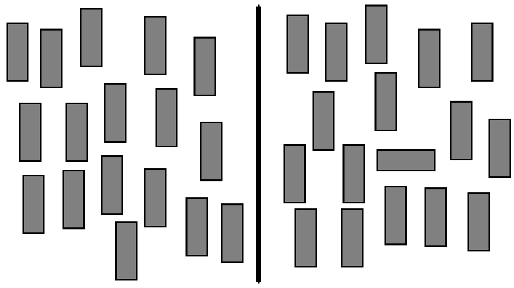


Figure 3.2: A search task where the target is defined by a single feature, orientation. On the left is a target absent trial, on the right is a target present trial - the target is the horizontal bar.

3.3.2 Features and Conjunctions

Of the early visual search experiments, some of the most influential were those conducted by Treisman and Gelade (1980). They produced two qualitatively different patterns of data from two different search tasks. They found that a flat search slope was produced when the target differed from homogenous distractors on a single featural dimension. In other words, additional distractors do not slow or inhibit the search process when the target 'pops out' from the background of distractors. An example of an easy visual search task that would produce a flat *set size* vs. *RT* graph is shown in Figure 3.2. This task requires the observer to look for a target that is defined against the background of distractors on one visual dimension (orientation: horizontal against a background of vertical, in Figure 3.2). Treisman and Gelade (1980) termed these easy searches *feature* searches, as the target could be discriminated from the distractors using information from a single featural dimension (orientation in Figure 3.2).

Treisman and Gelade (1980) found *set size vs. RT* functions with a slope were created for search tasks in which information from two featural dimensions was needed to discriminate the target from the distractors. An example of a difficult search task that would produce a steep *set size* vs. *RT* graph is shown in Figure 3.3. This task requires the observer to look for a target defined against the background on two visual dimensions - *orientation* and *colour* (or contrast).

In this instance, the target is a horizontal white bar in a background of bars that can be either horizontal or vertical and white or gray, but none of which are horizontal and white. Therefore, the observer must look for a target defined by a *conjunction* of the visual features present in the distractors[1].

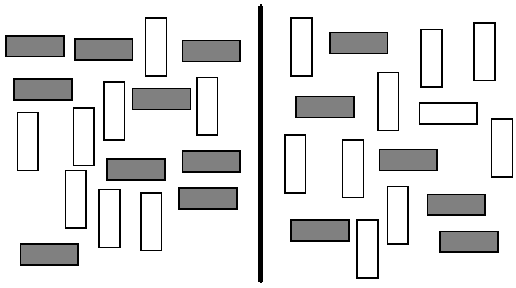


Figure 3.3: A search task where the target is defined by a conjunction of features, orientation and colour. On the left is a target absent trial, on the right is a target present trial - the target is the horizontal white bar.

Figure 3.4 shows results from target present and target absent trials for conjunction and feature (disjunction) searches. The feature searches were defined by colour or shape (e.g., a red target with blue distractors or a square target with triangular distractors) and the conjunction search was defined by a conjunction of colour and shape (e.g., a red triangle with blue triangles, red squares and blue squares as distractors). For both the conjunction and feature searches, the slope from target present trials is half that of the slope from target absent trials. The authors suggest this is because on target present trials, on average half of the elements must be searched before the target is found, whereas on target absent trials all the elements must be searched before it is decided there is no target. *Self-terminating search* refers to when the search stops on finding the target and target-absent slopes being twice target-present slopes is characteristic of this phenomenon (Treisman & Gelade, 1980). The other important result from this graph is that the target-present feature search trials create a slope of zero, while the target-present conjunction search trials create a definite positive

[1]The term 'feature' here is used to identify a particular point along a visual dimension or scale. As an example, orientation and colour are visual dimensions, horizontal and vertical are points on the dimension 'orientation' while white and gray are points on the dimension 'colour'.

slope. This pattern of results formed the basis of a theoretical framework of visual search, promoted by Anne Treisman and colleagues, known as Feature Integration Theory (FIT).

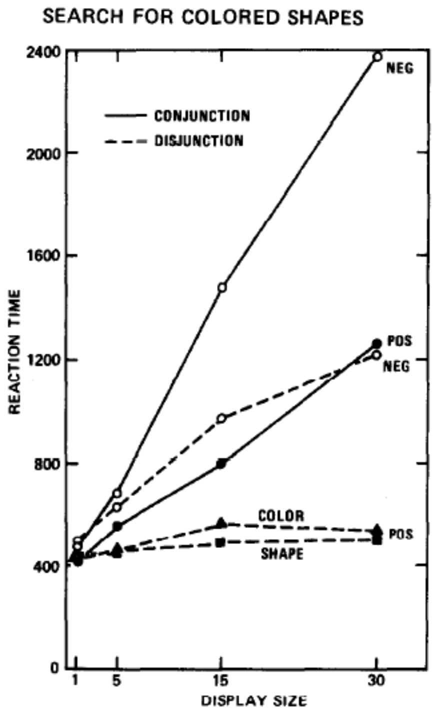


Figure 3.4: Data from the first experiment of Treisman and Gelade (1980). Slopes for positive (target present) and negative (target absent) trials are shown for both conjunction and disjunction (feature) conditions.

3.3.3 Feature Integration Theory

Feature Integration Theory (FIT, see Treisman & Gelade, 1980; Treisman, 1998) is schematized in Figure 3.5. The basic idea is that perceptual processing occurs in independent *feature maps* or *feature modules* which each represent the values on one featural dimension across the visual field. It is proposed that these feature maps exist for basic visual dimensions such as colour, size and orientation. FIT suggests that the outputs of the various feature maps are combined in a spatial location 'master map' and that this combination *requires attention.* However, this feature combination occurs only when it is needed by the search task. Therefore, when a stimulus can be differentiated from its background by the output of a single feature map, attention is not required according to FIT (this situation occurs in the *feature searches* discussed above, see Figure 3.2). Attention is, however, required to identify stimuli that can only be distinguished from their background by a conjunction of features. In other words, attention is required to identify stimuli that can only be distinguished by the output of several feature maps. This combination of outputs from several feature maps is termed ‘feature integration’ and is supposed to require attention, in contrast with detection based purely on ‘pop-out’ or ‘feature salience’[2]. In summary, FIT proposes two stages of processing for visual search tasks, a feature map stage and a feature integration stage. If the target cannot be isolated by the first stage, the second stage is invoked.

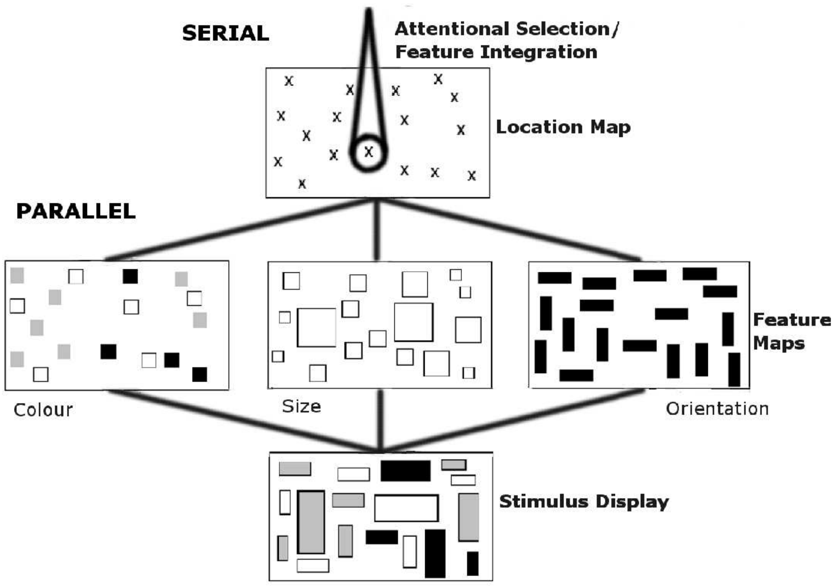


Figure 3.5: A depiction of Treisman’s Feature Integration Theory (FIT). Information from individual feature maps is combined spatially into a master map of locations. This combination requires attention.

The first stage in FIT is supposed to act in parallel across the visual field, meaning that each location in space is processed simultaneously and independently of other locations. This means this stage has no capacity limits, in the

[2]A salient stimulus can be considered as one which differs from its background on a single visual dimension (colour, orientation etc.) and has a background that is relatively homogeneous in that dimension (Humphreys & Bruce, 1989).

sense described previously (see Section 3.2). The second stage is supposed to act in a serial manner, meaning that only one location can be processed at a time. This stage therefore operates with a limited capacity. It is this parallel vs. serial operation that Treisman and Gelade (1980) used to explain the search slopes obtained by feature and conjunction tasks, respectively. Before FIT, the idea of parallel and serial processes operating depending on the nature of the task had been suggested by Neisser (1967), but this was in a more general context. FIT applied the idea to visual search situations specifically.

3.3.4 Guided Search Theory

Wolfe, Cave, and Franzel (1989) created a model of visual search called 'Guided Search' that, like FIT, uses a two stage parallel-serial framework. The main difference between Guided Search and FIT is that Guided Search supposes that parallel search processes can *guide* serial search processes by giving them access to the result of a pre-attentive, parallel search. For instance, if an observer is searching for a blue square amongst blue and red circles and squares (but no other blue squares), the parallel process could select all blue items and pass this information to the serial process which could then scan the blue items for a square. This would make the serial process more efficient. However, the efficiency of the serial process will vary depending on the ease with which the parallel process can select potential targets based on the guiding feature. This means that guided search predicts not a fixed slope for conjunction tasks, but a range of slopes depending on the features involved. Based on a review of search data, Wolfe and Horowitz (2004) concluded that colour, motion, orientation and size are undoubtedly *guiding attributes* in the context of guided search theory, while luminance onset, luminance polarity, shape and curvature are probable candidates to be guiding features.

3.3.5 Problems with FIT

3.3.5.1 Set-Size Effects and Search Efficiency

Visual search experiments conducted since the original formulation of FIT have shown that search tasks cannot be easily classified as serial- or parallel-processing tasks on the basis of search slopes (RT vs. set size). For instance, E. C. Carter and Carter (1981), R. C. Carter (1982) and Nagy and Sanchez (1990) gave observers a feature search task with a target defined by colour and found set size effects (i.e., an increasing RT with set size) when the colour of the target and distractors was close. In a similar experiment, Bergen and Julesz (1983) required observers to detect the presence or absence of a vertical line amongst a field of tilted lines, and modified the tilt of the distractors. For a given level of performance, the amount of tilt was inversely related to the display duration, meaning that a greater target-distractor difference was required for a shorter display duration. A single, continuous function relating orientation difference to performance was obtained in this study (see Figure 3.6a). This is unlikely to reflect the operation of two distinct processes - a serial process acting at low levels of discriminability and a parallel one acting at high levels. Therefore, it would appear from this data that performance for any given search task lies on a continuum related to the target-distractor discriminability, and that FIT's notion of two distinct search tasks is incorrect.

To more formally test whether data obtained from such experiments goes against FIT, Verghese and Nakayama (1994) used a limited capacity model as a basis for comparison. They tested search performance for orientation and colour and compared their data as well as the data from Bergen and Julesz (1983) (see Figure 3.6) with the predictions made by the *fixed sample size model of limited capacity* (Taylor & Creelman, 1967; Shaw, 1980; Palmer, Ames, & Lindsey, 1993)[3]

[3]This model assumes the observer makes a fixed number of samples in a given time, and so the number of samples is proportional to the duration of the display. With duration held constant, the model predicts target-distractor difference will increase with a log-log slope of -2

, and found that performance was better than that predicted by this model for all stimuli, except a limited range of spatial frequencies. Again, this evidence is in conflict with FIT, as it suggests a limited capacity (serial) process is not operating in the detection of any targets - even those with a low discriminability.

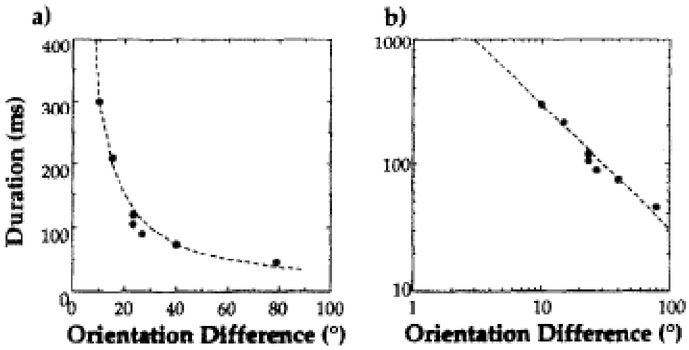


Figure 3.6: The data of Bergen and Julesz (1983) as originally presented (a) and as replotted in log-log space (b) by Verghese and Nakayama (1994).

The idea that set size vs. RT slopes lie on a continuum rather than being divided into two catgeories (serial and parallel) was also promoted by Wolfe (1998), who conducted a meta-analysis of 2,500 experimental sessions of visual search that had been carried out over 10 years in his lab. Figure 3.7 shows the distribution of search slopes for the entire data-set and demonstrates that this distribution is unimodal. Wolfe concludes that search slopes should be classified on a continuum, rather than into two or more categories based on this unimodality. However, Haslam, Porter, and Rothschild (2001) suggest that this data could be the result of two distinct processes despite being unimodal, because their statistical modeling techniques show that two latent distributions can combine or mix to form a single, skewed distribution (as found in the Wolfe study).

3.3.5.2 Capacity Limits and Search Slopes

Another problem with the original incarnation of FIT was in its assertion that capacity limits were necessarily implied by search slopes. Initially, it seemed

or more (i.e., a steeper slope) as a function of the number of elements. A shallower slope means the observer was not subject to limited capacity (Verghese & Nakayama, 1994).

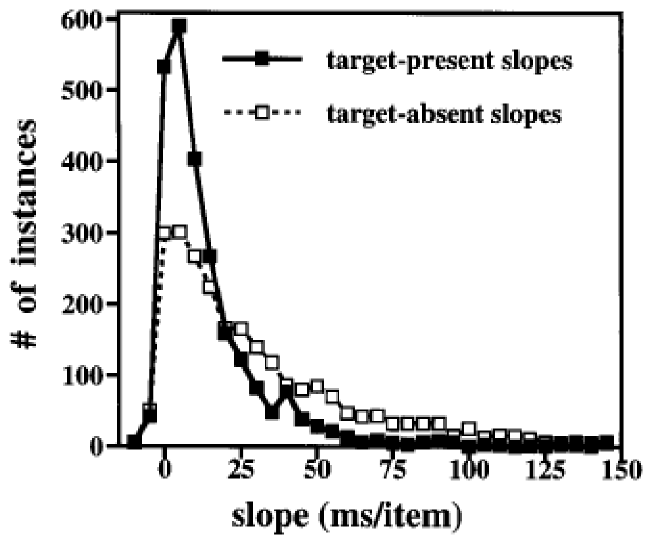


Figure 3.7: The distribution of results for 2,500 studies conducted in the lab of Jeremy Wolfe.

that attentional capacity limits can be inferred from search efficiency (given by accuracy vs. set size and RT vs. set size graphs). However, search time will increase and accuracy will decrease with set size even when there are no capacity limits due to the phenomenon of *statistical decision noise.* Statistical decision noise is present in any task when there is a non-zero probability that a distractor will be mistaken for a target (i.e., a false alarm). The probability of a false alarm occuring will increase with each additional distractor, and this decreases search accuracy (H. E. Pashler, 1998).

Huang and Pashler (2005) suggest that search efficiency - reflected by search slopes - reflects the task difficulty, but does not say anything about the involvement of attentional capacity limits. Essentially, the argument is based on the idea that an attentional capacity limit relates whether the quality or speed of processing of each individual item is affected by the presence of other item (i.e., does the quality/speed of processing decrease when items are added to the display). However, set size effects, they argue, can be affected by at least two other factors - statistical decision noise and eye movements. Therefore, attentional capacity limits can be considered independently of search efficiency (and search slopes).

What, then, is a method of measuring attentional capacity limits in a search task accurately? The task designed by Huang and Pashler (2005) compares a single presentation of a search field with successive presentations of parts of the search field. This is based on earlier studies by Eriksen and Spencer (1969) and Schriffin and Gardner (1972). In this task,

> "every element is presented for the same period of time and followed by a mask, so that the time available for processing each item is equated across the SUCC and SIM conditions" (Huang & Pashler, 2005) (SUCC = successive presentation, SIM = simultaneous presentation)

The idea is that if the task is subject to attentional capacity limits there should be a substantial advantage in the successive presentation condition as resources can be allocated to a subset of the display at a given time. Conversely, if there are no capacity limits, there should be no difference in performance between the successive and simultaneous presentation conditions. This is what the Huang and Pashler (2005) study found, despite there being substantial set size effects in both conditions. This result suggests that search performance should be conceptualised independently of capacity limits when theorising about visual search. A theory that does this is similarity theory.

3.3.6 Similarity Theory

A good framework for describing search results without involving theoretical notions such as serial and parallel processing is given in *similarity theory*, proposed by Quinlan and Humphreys (1987). This is based on the idea that search time and/or search difficulty depends on how easily a target can be distinguished from its background and also on the number of items of information required to identify the target. More formally, the model is described by the following principles (where 'T' refers to target and 'N' refers to non-target):

- as T-N similarity increases, search efficiency decreases and search time increases

- as N-N similarity decreases, search efficiency decreases and search time increases

- T-N similarity and N-N similarity are related; decreasing N-N similarity has little effect if T-N similarity is low; increasing T-N similarity has little effect if N-N similarity is high.

The use of the term *efficiency* by Quinlan and Humphreys (1987) to describe search performance is also suggested by Wolfe (1998) as a means of avoiding an implied theoretical position when describing search performance (i.e., as opposed to describing performance as 'serial' or 'parallel'). The description of search tasks in terms of efficiency without additional theoretical assertions is also utilised in the application of signal detection theory to visual search tasks.

3.4 Signal Detection Theory and Visual Search

3.4.1 High- and Low-Threshold Theories

The idea that search tasks may not easily be partitioned into parallel vs. serial processing has caused some researchers to consider visual search tasks as signal detection tasks, where a signal (target) must be selected from a background of noise (distractors). When the discrimination process sometimes fails to detect a target but will never mistake a distractor for a target, it has a *high threshold* for detection. Conversely, when it sometimes mistakes a distractor for a target but never fails to detect a target, it has a *low threshold* for detection. The threshold refers to the amount of signal information required for the presence of a target to be decided.

FIT can be considered a high threshold theory, because it does not allow for the influence of noise (e.g., statistical decision noise, mentioned above) on search performance. FIT invokes the serial-processing stage to explain search slopes. However, search slopes arise naturally from low threshold theories, through the accumulation of statistical decision noise (see Section 3.3.5.2). Therefore, instead of having both parallel and serial processing stages, signal detection models applied to visual search consist of a parallel processing stage followed by a decision rule [4] (Verghese, 2001). Furthermore, signal detection models do not have assumptions about limited capacity. Signal detection models of visual search can be considered low threshold theories, because noise is a central factor in their formulation.

3.4.2 Modeling Decision Stages in Search Tasks

Palmer et al. (1993) suggests one goal of the original formulation of signal detection theory (e.g., Green & Swets, 1966) was to distinguish between perceptual phenomena that affected individual percepts and decision phenomena that affect the mapping of percepts into responses. In signal detection theory, percepts are noisy and decision processes are necessary to map these into appropriate, task-relevant responses (Palmer, 1995). The noisiness of percepts predicts there will be set-size effects in visual search tasks even though the sources of information (the percepts) are independent - this is the statistical decision noise discussed previously. To model decision processes, assumptions must be made about the noise in perception and the rule for decision (Palmer et al., 1993). The decision model of set-size effects was initially developed by Green and Swets (1966) and Tanner (1961). It assumes that inputs to the decision are noisy. Signal detection theory supposes that when a stimulus is observed an internal, noisy representation is created that can be depicted as a normal distribution with a mean and standard deviation. The maximum rule for decision states that the observer indi-

[4] The decision rule does not involve processing of sensory input, but simply compares the output of the previous stage to an internal criterion/threshold. Decision stages in models involve integrating information from noisy perceptual representations in a way that is task-relevant (Palmer, 1995)

cates the target is the one that produced the largest percept. This is commonly used in two-interval forced choice tasks, while the sum rule (in which the sum of the percepts is used) is more appropriate for tasks with multiple targets and/or yes-no tasks (Palmer, 1990). To make predictions about these particular decision rules, Palmer (1990) assumes that the target and distractor are representable as stochastically independent Gaussian distributions. If T and D are random variables for the targets and distractors respectively, and f(x), g(x) and F(x), G(x) are density and cumulative distributions for targets and distractors, respectively, c is a decision criterion value, s is the stimulus value and w is a proportionality constant then it follows that the false alarm rate and the hit rate are given, for a set size of 1, as:

$$FA = F(D > c) = 1 - F(c) \tag{3.1}$$

$$H = F(T > c) = 1 - F(c - ws) \tag{3.2}$$

and for a set size of n, as:

$$FA = 1 - F(c)^n \tag{3.3}$$

$$H = 1 - F(c - ws)F(c)^{n-1} \tag{3.4}$$

3.4.3 Palmer's Studies

John Palmer has run a series of search experiments using a signal detection framework (see Palmer et al. (2000) for a summary). These experiments apply the methods of psychophysics to search in order to move toward a common account of search tasks and simple detection and discrimination tasks. To use the methods of simple detection studies, the visual search paradigm must be simplified and so Palmer et al. (2000) conducted search experiments with the following constraints, which are used in most simple detection studies:

- unidimensional judgements - judgements are made involving one stimulus dimension only
- accuracy, rather than reaction time, measures
- single eye fixation - presentation conditions are controlled so that participants must fixate in one position for the duration of a single trial presentation
- distinct and independent stimuli - stimuli are well separated and dispersed

These measures are all attempts to reduce the sources of noise in the task. Unidimensional judgements stop other dimensions competing for processing resources, or interfering with observer's judgements on a given dimension. Using accuracy measures means that stimuli are presented for a fixed time in each trial, eliminating another potential source of variation. Using only a single eye fixation means that the eccentricity of stimuli in the visual field is kept constant, eliminating variability due to differences in eccentricity, such as decreased perceived contrast, hue etc.. Using distinct and independent stimuli stops elements in the search field being automatically grouped, preventing configural cues that would affect task performance. All these constraints serve to reduce and control variability in the task.

3.4.3.1 Set-Size Effects Revisited

In relation to the signal detection approach to visual search, Palmer et al. (2000) states:

> "The predicted magnitude of a set-size effect depends critically on whether the theory assumes a high or low threshold".

Specifically, a high threshold theory predicts varying set size will have no effect, while a low threshold theory predicts there will be one.

Furthermore, the magnitude of the effect predicted by a low threshold theory depends on the discriminability of the stimuli. With more discriminable targets, the set size effect will be smaller (as predicted by similarity theory - see Quinlan & Humphreys, 1987). Therefore, it is necessary to equate stimulus discriminability across experiments to compare set size effects.

3.4.3.2 Difference Thresholds

Palmer et al. (2000) suggests the use of a difference threshold to equate stimulus discriminability across experiments. Figure 3.8 shows the difference thresholds plotted as a function of set size for individual observers and the mean of four observers for a contrast increment experiment conducted by Palmer (1994). These thresholds were calculated from psychometric functions plotted for each observer. The mean slope on this log-log plot is 0.30 ± 0.08. This slope can be used to compare the set size effect obtained in this experiment with those from other experiments, as the log-log transformation removes the dependence on scale. For this experiment in combination with the others conducted by Palmer (1994), the mean log-log set size slopes lie between 0.25 and 0.30. Such slopes can also be compared against the predictions made by various models of search behaviour.

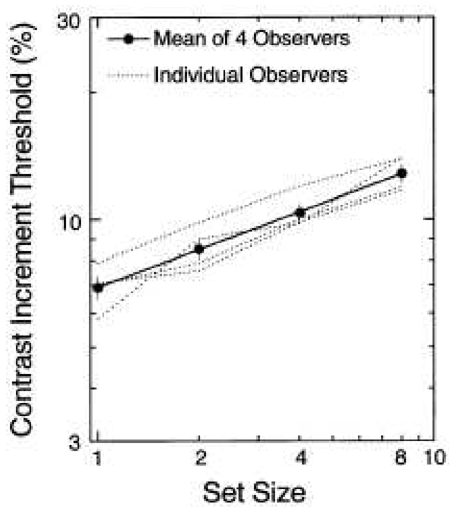


Figure 3.8: Thresholds for each of four observers and the mean of these observers from a contrast detection experiment by Palmer (1994). Taken from Palmer et al. (2000).

Palmer et al. (2000) ran simulations to determine the set size effects of various theories of search behaviour (see Figure 3.9). These simulations were based on Equations 3.2 and 3.4, but require some extra development, as follows. At threshold level, FA = 0.25 and H = 0.75. Therefore,

$$.25 = 1 - F(c)^n \tag{3.5}$$

$$.75 = 1 - F(c - ws(n))F(c)^{n-1} \tag{3.6}$$

Solving and simplifying (see the Appendix of Palmer et al. (1993) for more detail), gives:

$$s(n) = F^{-1}((.75)^{1/n}) - F^{-1}[(.75)^{1/n}/3] \tag{3.7}$$

w is a proportionality constant and the only free parameter in the equation. It can be assumed to be 1 as the predicted slope in log-log space is unaffected by it. For a yes-no task of line-length discrimination (Palmer et al., 1993), the predicted slope was 0.25. The high-threshold theory predicts a slope of zero while the other, low-threshold theories predict slopes in the range of 0.19 to 0.26 at a set size of 2, which is close to the obtained slopes. Similar set size effects to those obtained in the contrast increment experiment have been obtained for experiments using disk colour, line length, line orientation, rectangle shape, vernier acuity judgements, detection of Landolt Cs and letter discrimination (Palmer et al., 2000). However, tasks that require judging the orientation of pairs of widely separated objects and also tasks that make explicit memory demands have higher set size effects (Palmer et al., 2000). Palmer et al. (2000) suggests this is because these tasks are more complex and so may be subject to specialised, limited capacity processes that the other tasks are not.

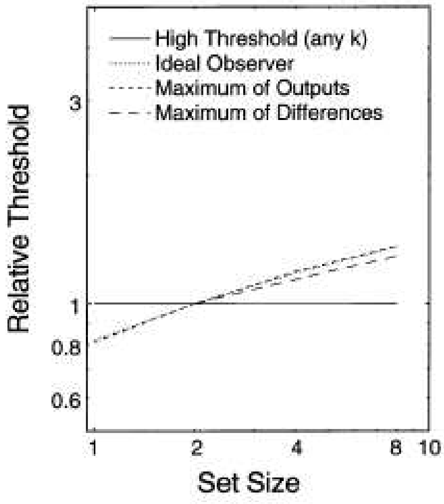


Figure 3.9: Model predictions for set size effects taken from Palmer et al. (2000).

3.5 Serial and Parallel Processing Revisited

The distinction between a positive RT vs. set size slope reflecting a serial process and a flat slope reflecting a parallel process is erroneous, as limited capacity parallel processing models will produce a similar straight-line slope to these serial models (Townsend, 1990). Linear increasing RT vs. set size curves can no longer be considered a fundamental parallel vs. serial distinction because it indicates simply that perhaps a limited capacity process is acting (Townsend, 1990; Huang & Pashler, 2005). What then is a legitimate way to distinguish serial from parallel processes? The multiple targets paradigm offers a way of making this distinction when RT is used as the dependent variable.

3.5.1 Multiple Targets in Visual Search

van der Heijden (1975) ran an early study dealing with the effects of multiple targets on search performance. In this experiment 1, 2 or 3 letters were presented around fixation. Two different letters were used, and one was defined as the target. Observers responded by indicating whether a target was present or not. A decrease in response times was found for an increasing number of targets both when those targets were presented as part of a fixed number of elements (i.e., sometimes with distractors) and when those targets were presented by themselves (i.e., with no distractors). van der Heijden (1975) took the decreasing response times with number of targets result in the no distractors condition to be evidence of a parallel self-terminating process. The facilitation effect of multiple targets is taken as evidence for a parallel process because it indicates that each target facilitates the decision that there was a target in the display. This means that information in each one of multiple channels is facilitating information in other channels or that the information is somehow summing. Figure 3.10 shows examples of predicted RT slopes in multiple target search and compares them with RT slopes for single target searches at each set size. A redundancy gain is indicative of a parallel process (not a serial process) and whether a parallel process is capacity limited or capacity unlimited depends on the slope of its single target RT vs. set size function. As a process becomes less parallel and more serial then, one would expect the redundancy gain to get smaller. A main point of the graph is that it is difficult to distinguish between a serial process and a capacity limited parallel process on the basis of single target slopes alone. A redundancy gain has been found for targets defined by translation direction (Thornton & Gilden, 2001), word detection (Mullin & Egeth, 1989) and letter detection (van der Heijden, 1975) (as mentioned before).

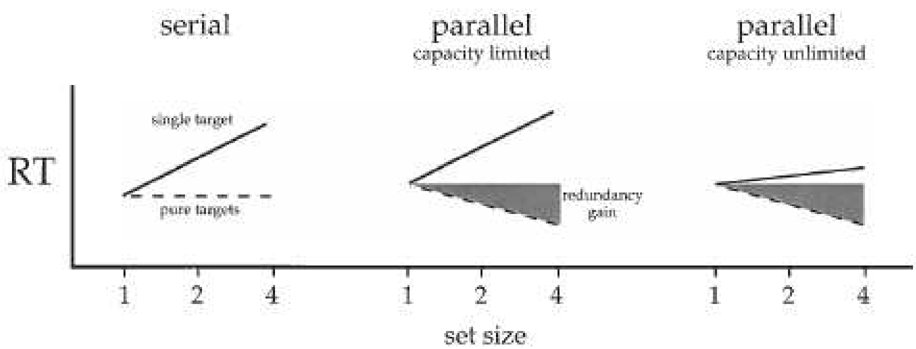


Figure 3.10: Patterns of predicted RT using in multiple-target searches, taken from Thornton and Gilden (2001).

3.6 The Medium of Visual Attention

The main distinction between early- and late-selection theories discussed previously (see Section 3.2) is where in the processing stream the selective function of attention operates. A similar distinction arises in relation to the level of representation on which attention acts. For instance, does attention select spatial locations or does it select integrated objects?

3.6.1 Space-Based Attention

Space-based theories of visual attention suggest that at any one moment, attention is focussed only on a given area of visual space and only stimuli within this area receive full perceptual analysis. This idea can be traced back to William James who suggested attention had a fringe or margin - in other words, that it was spatially limited (James, 1890/1950). A more modern conception is that,

> "Attention can be likened to a spotlight that enhances the efficiency of detection of events within its beam." (Posner, Snyder, & Davidson, 1980) p.172.

The spotlight account was originally based on experiments where observers were required to identify the location of a light in their visual field - this location was sometimes cued and sometimes not (Posner et al., 1980). A similar, but more elaborate, experiment was conducted by Eriksen and Murphy (1987), who presented a target adjacent to a distractor either 1, 2 or 3 degrees of visual angle away from fixation. In some instances, as in the experiments of Posner et al. (1980), there was a 'pre-cue' - a cross flashed at the location where the stimulus was going to appear. The distance between the target and the distractor was also varied trial-to-trial. For the distractors to have an effect (i.e., for reaction time to be slowed) in the pre-cue condition, distractors had to be placed close to the target and also had to be visually similar to the target. In the no pre-cue condition, distractors did not have to be close to the target or visually similar to it for them to have an effect.

These findings were integrated in a theoretical account that compared visual attention to a zoom lens, rather than a spot light. It was proposed that the resolving power and spatial extent share a complementary relationship such that an increase in one produces a decrease in the other (like a zoom lens) (Eriksen & Murphy, 1987). A pre-cue focuses the 'lens' in that region, decreasing the spatial extent while increasing the resolving power. If there is no pre-cue, however, the lens is unfocused by default and so covers a wide spatial extent but has low resolving power. The differences in resolving power between the pre-cue and no pre-cue conditions explain the effectiveness of the distractors.

The zoom lens account has been backed up by Shulman and Wilson (1987), who required observers to modify the contrast of the gratings until they could only just see them, while conducting a concurrent discrimination task using compound letter stimuli, which contain one letter at a large spatial scale and another at a smaller scale (see Figure 3.11). It was found that thresholds for finer gratings were lower when observers were making local discriminations, while thresholds for coarser gratings were lower when they were making global discriminations. According to the zoom lens account, the discrimination task would change the resolving power of the attentional lens. This would result in detection thresholds being better for stimuli matching the attentional task in terms of spatial frequency.

3.6.2 Object-Based Visual Attention

Several experiments have shown that attention can be distributed more efficiently to two aspects of a single object, rather than one aspect of each of two different objects. Treisman, Kahneman, and Burkell (1983) presented participants with a word and a rectangle and required them to read the word and locate a gap in the

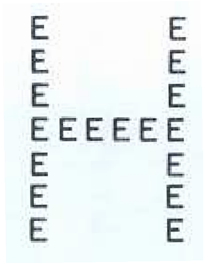


Figure 3.11: An example of a compound letter stimulus where the participant can be required to make a local discrimination (smaller letters) or a global discrimination (the larger, compound letter). Taken from Kinchla and Wolfe (1979).

rectangle. There were two conditions. In one, the word was inside the rectangle and in the other, the word was outside. Participants performed better when the word was inside the rectangle, even though the distance from fixation of the word and rectangle were the same in each case. A similar experiment was conducted by Duncan (1984), where two rectangles with gaps were presented and tilted lines went through each. Participants were required to make discriminations between the two objects or within one of them. Like the Treisman et al. (1983) experiment, within object discriminations were performed significantly quicker than between-object ones. This supports the idea that attention is selecting objects rather than spatial locations.

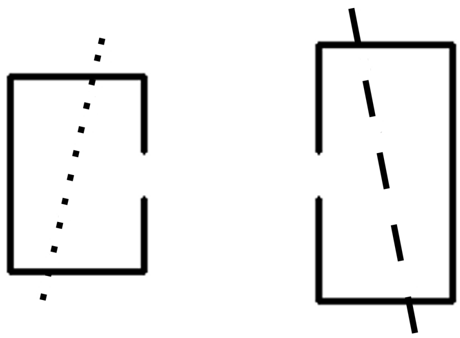


Figure 3.12: The stimuli used in the Duncan (1984) experiment, in which participants identified a single aspect of an object, two aspects of one object and one aspect of each of two objects. Identifying two aspects of one object was easier than identifying one aspect on each of two objects. Adapted from Duncan (1984)

More evidence for object-based attention came from a task in which participants were required to track a number of moving circles identified from a larger group of circles at the start of each trial (see Figure 3.13), and point to their location when they stopped at the end of the trial (Pylyshyn & Storm, 1988). This task is known as multiple object tracking (MOT), and a typical task will yield results showing that participants can reliably track around 4-5 objects that are each moving independently in random trajectories. A theory was constructed based on these results which suggests objects are pre-attentively localised and indexed, so that their location information is made available to other visual operations quickly and continually (FINST theory, see Pylyshyn, 1998).

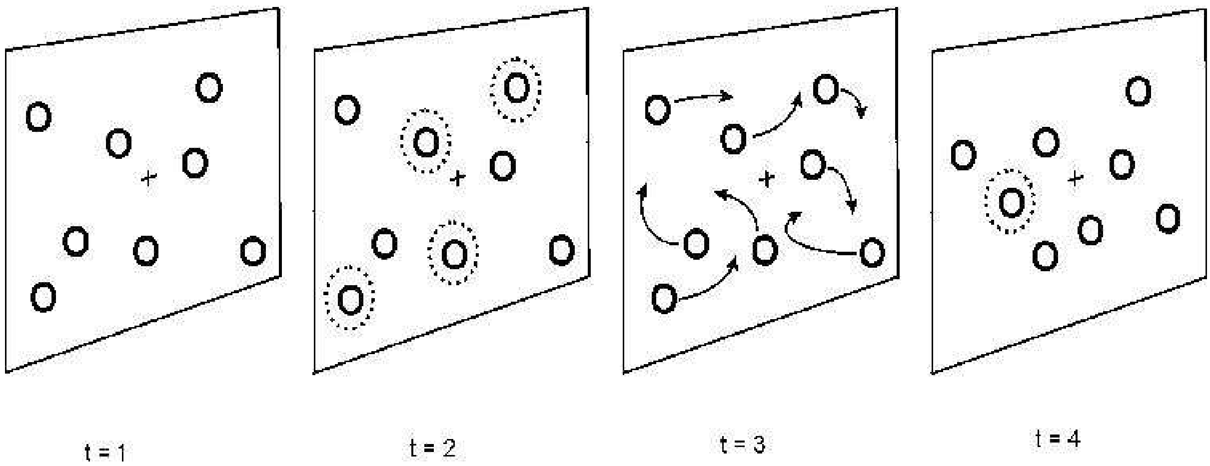


Figure 3.13: A schematic of a trial progression in a typical Multiple Object Tracking experiment. At t = 1, the elements are stationary; at t = 2, a subset of elements are each surrounded by a ring; at t = 3, all the elements move in random patterns; at t = 4, a single element is surrounded by a ring and the observer must decide whether or not this element was part of the original subset of elements identified by rings.

3.6.3 Multiple Mediums for Visual Attention

Evidence has emerged that both object- and location-based accounts of attention may be valid. Inhibition of return (IOR) is a phenomenon whereby the detection of targets at a previously attended spatial location is inhibited. A typical task demonstrating this phenomenon involves three boxes being presented on the screen, one in the centre, one on the right and one on the left. A cue is presented around the leftmost one and then around the centre one. After this, the target is presented inside the left (cued) or right (uncued) box and it is found that observers' reaction times are significantly longer when the target is in the cued position than in the uncued position (Maylor, 1985). When boxes other than the central one are moved during the trial, so that the target is presented on the same box but not in the same spatial position, reaction times to this box are still significantly slower than to others (Tipper, Driver, & Weaver, 1991). A modified task has shown that these spatial- and object-based IORs can exist simultaneously (Tipper, Weaver, Jerreat, & Burack, 1994). Therefore, the IOR paradigm demonstrates that the spatial- and object-based accounts of visual attention are not mutually exclusive but, instead, account for the patterns of performance found on different task types. Figure 3.14 shows both the Posner and Cohen (1984) and Tipper et al. (1991) tasks.

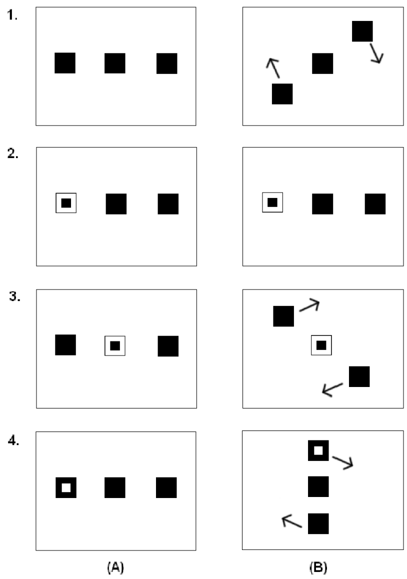


Figure 3.14: (A) shows the procedure of Posner and Cohen (1984), where the target (in 4) appears in a location that was previously cued (in 2). (B) shows the procedure of Tipper et al. (1991), in which the target (in 4) appears not in the same location as the cue (in 2), but on the same object. Reaction times to the target are slowed compared with equivalent no-cue conditions, demonstrating location-based and object-based IOR in (A) and (B) respectively. Adapted from Tipper and Weaver (1998).

3.7 What is selection for?

In this chapter, we have reviewed ideas on both the locus of visual selection and the medium or level at which it acts. Clearly, both issues are related. Another important question that is raised by such considerations is what is the information being selected for? We can invoke the notion of 'higher processes', but this is not a satisfactory answer, especially in the context of the current thesis in which we are looking at the retention of visual information over time. Therefore, in the next chapter, we look at what happens after visual information is selected and the theories that have arisen from such research. The next chapter focusses on visual memory.

Chapter 4

Visual Memory

Visual attention research, discussed in the last chapter, usually looks at behaviour and processes occuring across a number of fixations. Research on visual memory, however, began using methods presenting stimuli so briefly that only a single fixation could be made during their presentation. Stimuli presented for 100 msec or less are best for probing what happens in a single fixation, as the latency of an eye movement is always greater than 100 msec after a fixation is started (Coltheart, 1999). Despite the two different starting points, there are a number of issues for which research on visual attention and visual memory overlap, such as capacity and the nature of visual encoding.

4.1 Sperling's Experiments

Using stimuli presented on a tachistoscope, Sperling (1960) developed a paradigm for studying the nature of visual processes occurring within a single fixation. He presented subjects with a 3×3 or 3×4 array of letters for 50 msec and they were told to remember as many as possible. He found that subjects reported only four or five letters. However, after considering that this might reflect a limit on the number of letters subjects could *recall*, rather than the number they could *store*, he modified the experiment.

In the modified experiment, an auditory tone sounded after the offset of the visual display, where the pitch of this tone identified the row of letters the observer must report (observers had previously been trained with what tone matched what row of letters). Therefore, in each trial the observer is asked only to report a single row of three letters. In this experiment, observers successfully reported whatever row they were cued to report. However, observers were completely accurate (i.e., all three letters correct) only if the tone followed the visual stimulus by 100 msec or less. Accuracy decreased with longer delays (see Figure 4.1).

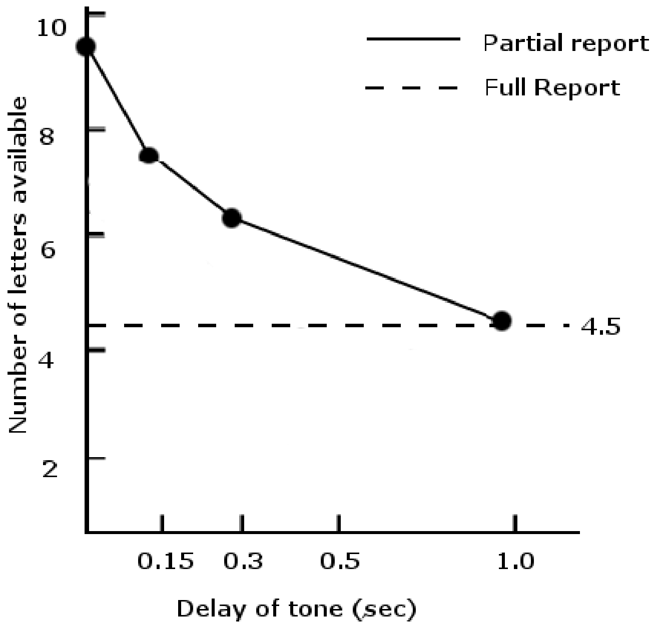


Figure 4.1: The results of Sperling's partial report method. The 'number of letters available' measure is calculated from the number correct out of the partial report.

Sperling (1960) also found the post-exposure field could affect the retention of information from the original display. When this field is black, observers can recall information for up to five seconds, whereas if it is white, they can recall it only up to half a second. This effect was later termed 'visual masking'. A stimulus functions as a *mask* if it reduces the detectability of the target stimulus (Turvey, 1973; Coltheart, 1999). A mask presented before the target stimulus is termed a 'forward mask' and one presented after is termed a 'backward mask', indicating the position they have relative to the target stimulus and the direction (in temporal terms) of their influence upon memory[1]. Usually, a mask is not a

[1]Additionally, experiments can use 'lateral masks' which are presented in the same display as

blank field but a scrambled spatial pattern that, as well as masking the to-be-remembered stimulus, ensures retinal afterimages from it do not persist.

Averbach and Coriell (1961) conducted a study similar to that of Sperling (1960), in which an array of letters was presented for 50 msec, followed by a blank white field for a variable amount of time, followed by a partial report cue (either a bar above one of the positions previously occupied by a letter, or a circle surrounding it) for 50 msec, followed by another blank field. Participants were required to report the letter identified by the partial report cue. Results for this study are shown in Figure 4.2. When the bar is used as the partial report cue, performance decreases steadily over the first 100 msec and then reaches a plateau. When the circle is used, performance is greatly decreased in the period 50 msec - 150 msec. This is probably due to a masking effect, as the circle occupied the space of the number. Both the Sperling (1960) and Averbach and Coriell (1961) studies appear to indicate a high-capacity and fast-decaying form of visual memory that persists for around 100 - 150 msec. Neisser (1967) called it *iconic memory* and suggests it is a perceptual form of memory, as observers report they 'see' the original stimulus during recall.

It is clear from these early studies that information from a single fixation is represented in detail by a system that has a high capacity and decays quickly. This begs the question of whether there is a memory acting *across* fixations and, if there is, what is the nature of it. To answer this question, we must first return to issues associated with the movement of eyes across a visual scene.

4.2 Integration of Information Across Fixations

In Chapter 2, Section 2.2.1, the need for eye position information in scene perception was discussed, as was physiological evidence of a mechanism that obtains this

the to-be-remembered stimulus and are most effective when they are presented close to it. These masks are spatially separate from the target stimulus, but not temporally separate. Therefore, masking appears to act both spatially and temporally under the right conditions.

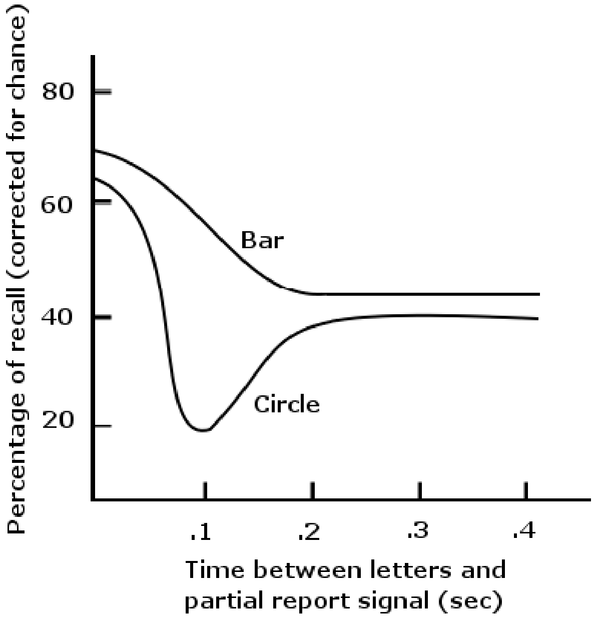


Figure 4.2: The critical results of the Averbach and Coriell (1961) study, for the conditions where a bar or a mask was used next to/surrounding the target letter.

information from the eye muscles. From this evidence, it appears likely that eye position information is obtained from efferent motor commands and this information is somehow used to facilitate stability of the visual world. A theory that attempts to further explain how stability is maintained across eye movements is called the *spatiotopic fusion hypothesis* by Irwin (1993), which he states as:

> "the visible contents of successive fixations are spatially reconciled across changes in eye position and superimposed according to environmental coordinates at an early stage of perceptual processing" (Irwin, 1993) p.122

Irwin (1993) suggests this hypothesis has been promoted in various forms in the literature, in different amounts of detail. For instance, McConkie and Rayner (1976) suggested that the contents of each fixation is spatially integrated with the contents of others in an *integrative visual buffer*. Several experiments that appeared to support the spatiotopic fusion hypothesis (e.g., Ritter, 1976; Wolf, Hauske, & Lupp, 1978) were later shown to have methodological problems (see Irwin, Zacks, & Brown, 1990; Irwin, Yantis, & Jonides, 1983) that compromise the interpretation of their data. In addition, other experiments have provided evidence against the existence of a highly detailed memory of the spatial layout of the environment.

Irwin (1991) examined the spatiotopic fusion hypothesis with a task in which observers were required to detect if two dot patterns were the same or different. Experiment 1 contained three conditions. In one, the two patterns were presented the same distance from fixation, in the same spatial location (retinal and spatial overlap). In another, one pattern was presented left of fixation while the other was presented above (no overlap). In the third, both patterns were in the same spatial location, but the second appeared only after the observer had made a saccade to a second fixation point - therefore the patterns were not in the same retinal location (saccade condition - spatial overlap, no retinal overlap). Figure 4.3 shows performance for the conditions across different ISIs. For ISIs of less than 500 msec, performance declined for the retinal and spatial overlap condition and increased for the other two conditions. This suggests performance in the first condition relied on a high-capacity, fast-decaying memory while performance in the other conditions did not.

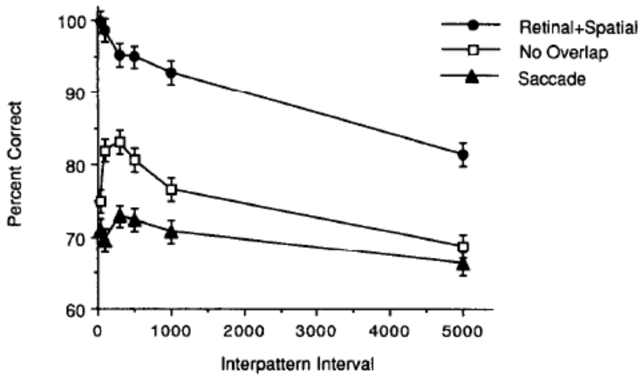


Figure 4.3: Performance for the three conditions in Experiment 1 of Irwin (1991), in which the spatial and retinal correspondence of the two stimuli to be compared were manipulated.

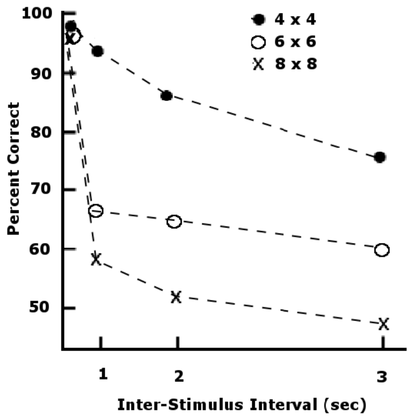
Given that performance across a saccade did not appear to use a high-capacity memory, Irwin (1991) tested whether performance across saccades was sensitive to the displacement of the pattern. If an iconic-like store were being used to retain information across saccades, it would be highly sensitive to displacements of the pattern. In the first condition, the post-saccade pattern was presented in the same location, relative to fixation, as the pre-saccade pattern. In the second condition, the post-saccade pattern was displaced slightly. It was found that this displacement did not affect observers' same/different accuracy.

Because it is not sensitive to pattern displacement, it appears that the memory acting across saccades does not reconcile spatial information in an exact, metric manner. Neither does it have access to a high-capacity store, like iconic memory. Irwin (1991) suggests instead that it is a limited-capacity, schematic store that is general purpose in nature, rather than specific to saccade-based activities. A good candidate for such a system is Visual Short-Term Memory, a system described as having a capacity of around 4-5 visual objects that lasts for around 30 seconds (Baddeley, 2003).

4.3 Early Experiments on Visual Short-Term Memory

A short-term store for visual information was first described in relation to studies in which observers were required to determine whether two matrices (Phillips & Baddeley, 1971; Phillips, 1974) or figures (Cermack, 1971), separated by a short delay, were the same or different. Phillips (1974) used arrays of blocks in his task so that participants could not associate any verbal or semantic information with the stimuli. On 'different' trials, the second array was changed from the first by the addition or deletion of a single block. In Experiment 1, the size of the matrix and the inter-stimulus interval (ISI) were manipulated and it was found that using larger, more complex patterns resulted in decreased performance, but only for ISIs longer than 500 msec. In Experiment 2, another manipulation was introduced - on half of the trials, the second pattern was displaced relative to the first and this decreased performance for ISIs less than 500 msec. In Experiment 3, a masking pattern (a 10×10 array of blocks) was flashed in the ISI of half of the trials, and this too decreased performance for all ISIs. See Figure 4.4 for the results from Phillips' experiments.

These experiments demonstrate that, in the 500 msec following presentation, there is no effect of pattern complexity but there is an effect of pattern displacement and an effect of masking. After 500 msec, an effect of pattern complexity emerges, but the effects of displacement and masking disappear. Phillips interpreted this dissociation as reflecting the operation of two distinct visual memory stores at different time scales: a high-capacity, retinotopic sensory store that retained information for up to 500 msec and a limited-capacity short-term store that retains information in a more schematic form, for at least 10 seconds. Phillips suggested the sensory store was the iconic memory referred to by Neisser (1967) and termed the schematic store the short-term visual store (STVS).



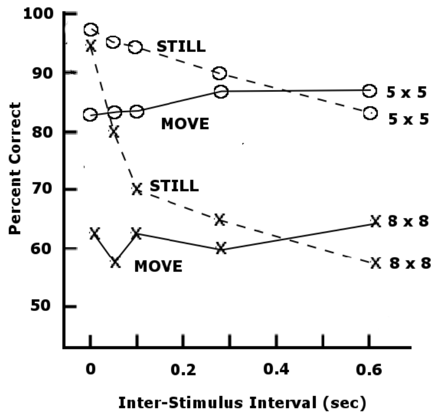


Figure 4.4: Left: Results from Phillips (1974) Experiment 1, showing how performance varies with ISI for different array sizes. Right: Results for Phillips (1974) Experiment 2, showing how performance varies with ISI, array size and displacement of the post-change array.

In addition to differentiating the STVS from iconic memory, Phillips ran experiments to differentiate it from long-term memory. When participants were required to remember a series of matrices, the most recent one was remembered far more accurately than any of the others, for which accuracy was fairly similar (Phillips & Christie, 1977). Furthermore, speeding up the rate of presentation decreased performance for all but the last matrix. Another experiment (Phillips, 1983) showed that performance on this task is disrupted when participants concurrently perform another visual task (one which does not require retention of information). These results are indicative of a short-term rather than a long-term store, as long-term memory typically contains more robust representations. The properties of the STVS revealed in these experiments are similar to the properties of verbal memory, which also shows a recency effect[2] (Murdoch, 1962) and disruption when irrelevant verbal information is presented (Salamé & Baddeley, 1982).

The experiments summarised above used complex stimuli and the information that needed to be remembered was relatively uncontrolled. This means that these experiments cannot answer more specific questions about vSTM. An example of a more specific question that has been explored more recently in the literature is: At what level does vSTM encode visual information - at the level of individual feature dimensions or at the higher level of integrated objects?

4.4 The vSTM Representation

Logically, there are several ways that vSTM could store information. For instance, Wilken (2001) suggests an 'item' in vSTM could represent the following:

- a coherent object made of multiple features
- a sub-component of an object, perhaps defined by common luminance or colour contours, itself made of multiple features

[2]A recency effect occurs when more recently presented stimuli are remembered better than previously presented stimuli.

- a single feature of an object

Another logical possibility is that an item could represent different things (e.g., an object or a feature) depending on the task during which encoding is taking place. The literature is inconclusive as to what an item in vSTM represents, but does include evidence for all of the aforementioned possibilities.

4.4.1 Feature-Based vSTM

To compare vSTM performance on different featural dimensions, several experiments (Magnussen et al., 1991; Magnussen & Greenlee, 1992, 1997) compared discrimination thresholds[3] across different time intervals, for different visual features (see Figure 4.5 for an example task). It was found that, with the exception of contrast, discrimination thresholds were almost exactly the same when the two stimuli were separated by several seconds compared to when they were presented simultaneously. Furthermore, another experiment found little change in spatial frequency thresholds over 2 days (Magnussen & Dyrnes, 1994). It was also found that a mask containing information from the same stimulus dimension decreased accuracy significantly, but one containing only information from another stimulus dimension did not. Therefore, Magnussen (2000) suggests the form of memory revealed in their experiments stores information from different stimulus dimensions in parallel, but does not allow the transfer of information between dimensions.

4.4.2 Chunking

There is also evidence for a visual short-term memory storing information as complete objects, rather than as features in parallel. This idea is related to the more general idea of *chunking* introduced by Miller (1956). Chunking refers to the process by which elementary bits of information are integrated or grouped into

[3] A discrimination threshold is obtained by presenting the observer with two or more stimuli, among which one is different to all the others, which are the same. The level of difference is manipulated across trials and the threshold is the difference required for a performance level of 75% to be attained.

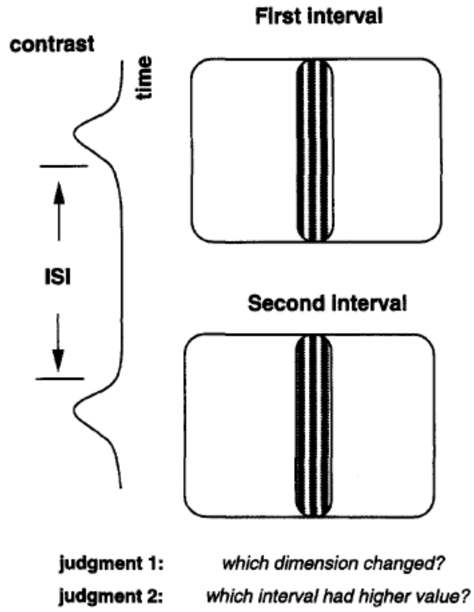


Figure 4.5: A schematic of a trial in the Magnussen et al. (1991) study, involving a change in contrast.

one chunk. The chunk then becomes the unit of storage for the memory system, such that its capacity is limited by the number of chunks, not the amount of elementary information. The idea that vSTM codes information as complete objects is analogous to the idea of chunking because it holds that the unit of storage (the object) is built from the integration of more elementary bits of information (features).

4.4.3 Evidence from Recall Tasks

Allport (1971) presented observers with three coloured shapes followed by a blank screen and then a mask soon after. Reports of both colour and shape were found to be as good as reports of colour or shape alone. A similar pattern of results was obtained when observers were asked to report the orientation and spatial frequency of a patch of lines (Wing & Allport, 1972). Duncan (1984) asked observers to report one or more properties of a line and/or a rectangle which the line was passing through. He found an object-based benefit, that is observers could respond accurately reporting multiple features of the same object without a cost, but could not report multiple features belonging to different objects without a cost. Irwin and Andrews (1996) presented observers with an array of 6 or 10 coloured letters briefly and asked them to report both the identity and colour of a letter at a specified location. They found that participants could correctly report both properties of 3-4 items, the same number they could report in a task requiring them to only report the identity of the items (see Irwin, 1992). These results appear to support an object-based account of vSTM, where the capacity of vSTM is determined by the number of integrated objects that are selected, not by the number of features. More recent support has come from Luck and Vogel (1997).

4.4.4 The Luck and Vogel (1997) Study

Looking at the question of how vSTM codes information, Luck and Vogel (1997) conducted a study in which participants were required to monitor an array of elements across a blank interval, during which one of the elements could change on one or two features. It was found that performance decreased with the number of elements when this number was greater than four. Furthermore, the performance vs. number of elements function was the same when participants were required to monitor for one of two changes compared to when they were required to monitor for just one change (see Figure 4.6). Therefore, it appeared that it was the number of *objects* rather than the number of *features* that determined performance and so Luck and Vogel (1997) suggested vSTM codes information as complete objects, rather than separate features.

The Luck and Vogel (1997) results can also be explained by the featural information being stored in parallel channels, where performance for any given feature is dependent only on the amount of information stored in that featural channel and no other featural channels (Alvarez & Cavanagh, 2004). However, the study had a condition that caused them to reject the parallel feature module account. In

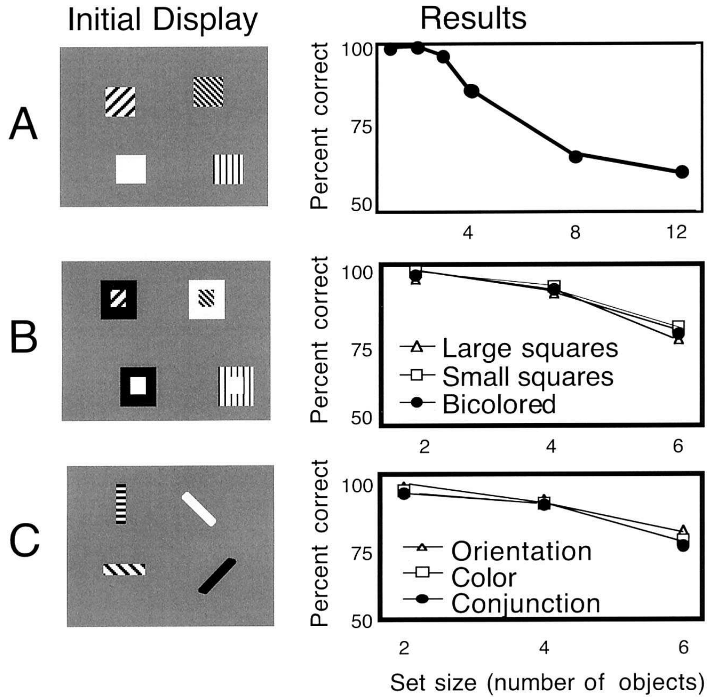


Figure 4.6: Stimuli and results for the critical conditions in the Luck and Vogel (1997) study. Different textures represent different colours in the actual stimulus display. A - unicoloured squares. B - bicoloured squares. C - orientation and colour. Taken from Wheeler and Treisman (2002).

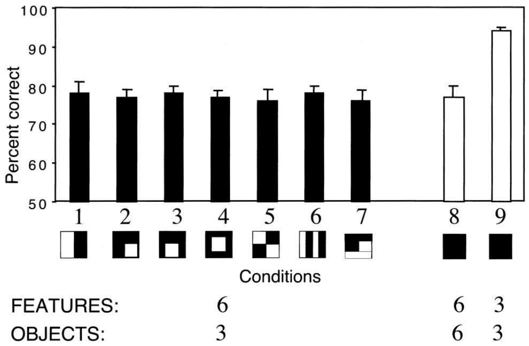
this condition, they presented observers with bicoloured squares - a large coloured square containing a smaller one of a different colour. They found performance for monitoring either colour (outside or inside) for a change was the same as when only one was monitored. Therefore even when two features in the same channel (colour) were to be monitored, performance did not decrease relative to the condition where only one is monitored. This undermines the feature-based interpretation, because it suggests that multiple features within the same channel are monitored in parallel when they belong to the same object.

4.4.5 Evidence Against the Critical Condition of Luck and Vogel (1997)

Experiments 1-4 of Patrick Wilken's PhD thesis (see Wilken, 2001), re-examined the findings of Luck and Vogel (1997) and failed to replicate the results of their critical bi-coloured condition. Specifically, Wilken (2001) found that performance decreased with the number of features changing, rather than with the number of objects changing, even though conditions in Experiments 1-4 got successively closer to those of Luck and Vogel (1997). However, the set size effects found by Wilken (2001) were similar to those of Luck and Vogel (1997). Wilken (2001) suggested that it was unclear why their findings differed from those of Luck and Vogel (1997), considering the similarity of the experimental conditions. However, the results of Wilken (2001) are somewhat in agreement with those of Wheeler and Treisman (2002).

Wheeler and Treisman (2002) also re-examined the findings by Luck and Vogel (1997) and also failed to replicate the results of their critical condition, in which bicoloured objects were used. Wheeler and Treisman (2002) used a range of bicoloured conditions and, in each, the two colours were used on the elements in a different configuration (see the x axis in Experiment 1 in Figure 4.7). In one of these conditions (Condition 4 in Figure 4.7 Left), the colours were configured in the same way as the Luck and Vogel (1997) experiment. Wheeler and Treisman (2002) found performance in these bicoloured conditions was the same as a control condition in which six unicoloured objects needed to be monitored, and significantly less than performance in a condition with three unicoloured objects (see Figure 4.7). Wheeler and Treisman (2002) conducted another experiment comparing performance for bicoloured squares with performance for an equal number of small unicoloured squares or an equal number of large unicoloured squares. Luck and Vogel (1997) found these unicoloured conditions produced the same level of performance as their bicoloured condition, but Wheeler and Treisman (2002) found performance was significantly lower for the bicoloured condition (see Figure 4.7 for the results from these experiments). Although this result would tend to favour a feature-based account, it does not provide a direct test of the object-based account.

The fourth experiment of Wheeler and Treisman (2002) more directly tested



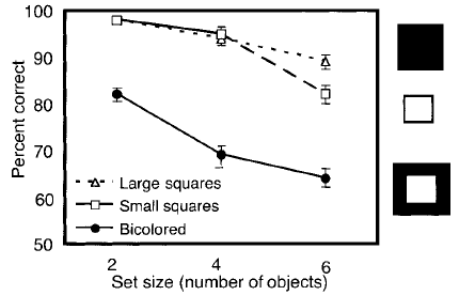


Figure 4.7: Left: Results of Experiment 1 from Wheeler and Treisman (2002), showing performance in a range of bicoloured conditions. Right: Results from Experiment 2 of Wheeler and Treisman (2002), showing performance in the unicoloured large and small squares conditions compared to performance in the bicoloured condition.

the idea that vSTM encodes information as complete objects. In this experiment, colour or shape could change from pre-test to the test display. The conditions in this experiment were as follows:

- Participants monitored items for a colour change, on *different* trials (i.e., trials where the post-blank display was different to the pre-blank one) two elements changed to colours at test that were not present at pre-test
- Participants monitored items for a shape change, on *different* trials two elements changed to shapes at test that were not present at pre-test
- Participants monitored items for a colour or shape change, on *different* trials at test, an element changed colour or shape
- Participants monitored for squares changing place with each other, on *different* trials two elements swapped colours (or two shapes swapped colours)

In the final condition, the relationship between colour and shape changed for two of the elements. This means the way these features were bound together in the display changed from pre- to post-blank. This condition provided a direct test of whether performance relied on features being bound into complete objects. If features are bound, the swapping of two elements should be equivalent to a change in one feature in terms of its effect on performance. However, performance in this condition was signficantly below that of the single-change condition. The authors re-ran these conditions using a modified post-blank presentation in which only a single element was presented (a *single probe* test). This was done because experiments yielding results in support of the integrated-feature approach (e.g., Ceraso, 1985; Irwin, 1992) had used partial report methodologies or recall, in which only a portion of the pre-blank display, or none of it, was presented to observers at test. Using this modified response paradigm, it was found that performance in the binding condition was the same as performance for the colour-only change condition. This is what would be expected if vSTM were acting at the level of integrated objects (with features bound together), rather than at the level of features. Therefore, the interpretation of the nature of vSTM encoding changes depending on whether more than just the changing element is present to participants at test.

Wheeler and Treisman (2002) suggested the poor performance observed in the binding condition for a whole-display response could be due to distraction by the other elements or, alternatively, a higher load being placed on decision making processes. To test these hypotheses, they ran another experiment in which the changing item was cued so that the participants only had to decide whether one item was changing. When the changing item was cued, performance was the same as when it wasn't. Therefore, the authors concluded that the selection of the changing item was being interfered with. More specifically, they suggested that *focussed attention* was being misdirected by the presence of distractors and that this process is required for integrating information into an object-based form, suitable for vSTM. The idea that attention is required to integrate features into objects has been proposed elsewhere (Treisman & Gelade, 1980; Kahneman, Treisman, & Gibbs, 1992), and was discussed in the last chapter. This idea of feature integration or feature binding is where discussions of attention and memory overlap. We could say attention selects and transforms the information into a form that memory can store. However, it is unclear whether it is the selective process that transforms the information, or whether this occurs in memory. In fact, this may be an unnecessary and arbitrary distinction to make. A paradigm that looks more closely at the relationship between visual attention and visual memory is that of *change blindness*, and this is discussed in the next chapter.

Chapter 5

Change Blindness

5.1 Induced Blindness

Chapter 3 discussed how information is selected from the visual field through the process of attention and Chapter 4 discussed how this information can be retained for later use. In this chapter a phenomenon called *change blindness* is examined. This phenomenon has been attributed to a failure of selection, a failure of retention or a failure of both (see Simons & Rensink, 2005). Change blindness refers to the difficulty observers have in detecting changes to a scene when those changes are obscured or made gradually, or when their attention is diverted from the change. Change blindness is a type of *induced blindness*. An *induced blindness* is any deficit in detection of an otherwise highly-visible stimulus that occurs when an observer's attention is somehow diverted from that stimulus.

5.2 Inattentional Blindness

5.2.1 Gorillas in the Midst

A type of induced blindness, known as inattentional blindness, occurs when observers attend to one stimulus and fail to report the presence and/or characteristics of another. The earliest example of this phenomenon is provided by Neisser and Becklen (1975), who used a mirror-based apparatus to superimpose multiple video sequences and present the composite sequence to an observer. In one experiment, one video sequence was of people in white shirts passing a ball amongst themselves and another was of people in black shirts doing the same. Observers were required to attend to one of the two teams and press a key whenever that team made a pass. After 30 seconds, another overlayed video sequence was displayed in which a woman carrying an umbrella walked across the display for 4 secs (see Figure 5.1). Only 21% of observers reported seeing the woman. In another version of the task, the woman wore the same coloured shirt as one of the teams, but this was found not to affect detection of her (Neisser, 1979).

Figure 5.1: The overlayed sequence of videos displayed to participants in the Neisser and Becklen (1975) study. The woman carrying the umbrella is in the centre of the frame and the umbrella is white.

Simons and Chabris (1999) conducted an experiment where the same events as the Neisser (1979) study occurred (i.e., teams passing balls, umbrella woman walking through) but these events occurred in the same video, rather than in separate overlayed videos. Simons and Chabris (1999) also included a condition in which the woman walked through the people passing the balls in a gorilla suit. The people in the video were presented as either transparent to the background, or opaque (see Figure 5.2 for the four different display conditions). Participants were required to monitor only one team - wearing either white or black shirts and were given either an easy monitoring task, in which they counted the total number of passes made by the attended team, or a hard monitoring task, in which they counted separately the number of bounce and aerial passes made by the attended team. Each participant participated in only one of 16 conditions (4 display conditions $\times$ 4 task conditions).

Figure 5.2: The four different display conditions in the Simons and Chabris (1999) study. For each display condition, there were four task conditions, where the team the participant attended to was manipulated (white or black) as was the task they had to perform (counting passes or counting bounce and aerial passes). In total then, there were sixteen conditions.

After performing the task, participants were asked to write their counts (of basketball passes) down and also provide answers to questions regarding if they saw anything unexpected. Out of all participants across all conditions, 54% noticed the unexpected event and 46% failed to notice it. More participants noticed it in the opaque conditions (67%) than in the transparent conditions (42%) and more noticed the unexpected event when performing the easy monitoring task (64%) than when performing the hard monitoring task (45%). These results can be easily explained by visibility and task load, respectively. More people noticed the umbrella woman (65%) than the gorilla (44%), which could be due to a difference in low-level salience. However, more people noticed the gorilla when they were attending to the black team than when they were attending to the white team. No such difference was found for conditions involving the umbrella woman. Therefore, what stimulus observers were attending to (i.e., white vs. black team) affected the rate at which they detected the gorilla, an unexpected stimulus. This was not found in the Neisser (1979) study. Simons and Chabris (1999) concluded that observers are more likely to notice unexpected events if these events are visually similar to those they are attending to (the umbrella woman was not similar to the white or black team, while the gorilla was similar to the black team). Logically, however, there are two possible reasons for this: similarity to the attended events and dissimilarity to the ignored events. To differentiate between such possibilities, it is necessary to use a more controlled experimental paradigm.

5.2.2 The Experiments of Mack and Rock (1998)

A more controlled series of inattentional blindness experiments was carried out by Mack and Rock (see Mack & Rock, 1998 for a review of these experiments). A typical task of theirs involves the participant judging which of the two arms of a fixation cross is longer, and on one trial of this task an unexpected object is presented at the same time as the cross (see Figure 5.3). At the end of a run of trials, participants are asked if they noticed anything unusual and about 25% do not detect the unexpected object. This result is the same regardless of whether this object is composed of a single feature salient on a dimension thought to be detected pre-attentively (e.g., colour, orientation, motion) or whether it is not. On the basis of such experiments, Mack and Rock (1998) proposed that observers do not consciously perceive events to which they do not attend.

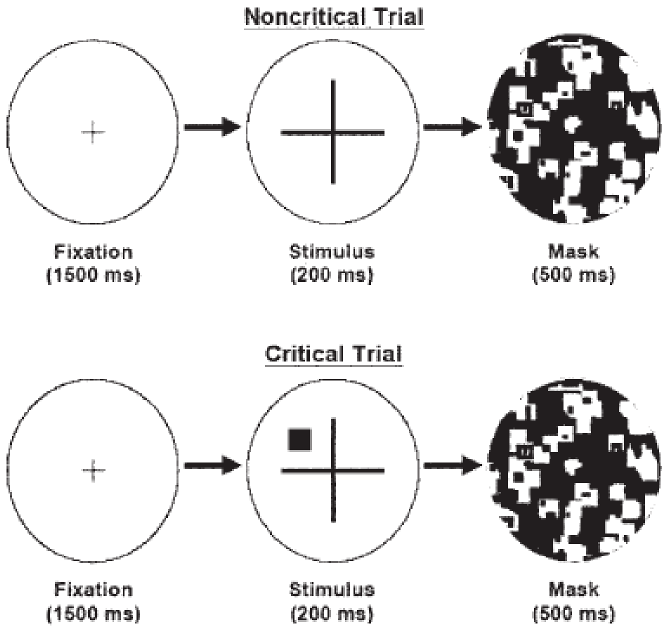


Figure 5.3: An example of a non-critical (no unexpected object) and a critical (unexpected object present) trial from a study described in Mack and Rock (1998). Participants are required to judge which arm of the cross is longer and are later questioned about whether they noticed anything unusual/unexpected during the trial.

5.2.3 The Experiments of Most et al. (2001)

Studies of inattentional blindness have been, until recently, limited to the use of static, unmoving stimuli. A few years ago, Most et al. (2001) used a presentation paradigm in which all objects, including the unexpected object, were moving all the time. The unexpected object was a cross and the other objects were letters, half of which were white and half of which were black. The cross was either black, white, light gray or dark gray (see Figure 5.4). In Experiment 1 of this study, the authors investigated whether the similarity of the unexpected object to the other objects influences its detection. Experiment 2 was designed to answer the question of whether the effects of similarity on inattentional blindness are driven by selective attention to the attented items or selective ignoring of the other items. Experiment 3 looked at whether inattentional blindness occurs at a higher rate when the unexpected object is different to the other objects on multiple featural dimensions.

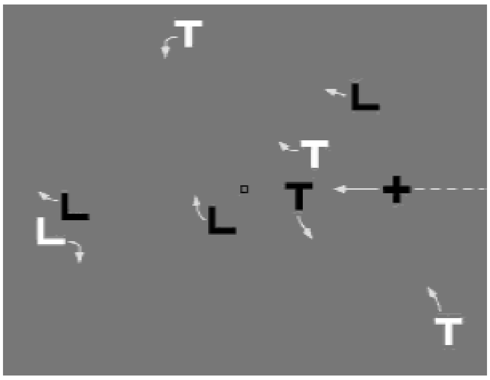


Figure 5.4: A depiction of trials in Experiment 1 of the Most et al. (2001) experiment, with arrows indicating the motion of elements and the dotted line indicating the path of the unexpected object (the cross).

Results from Experiment 1 showed that detection rates were highest for crosses of a colour/luminance most similar to the attended items and lowest for those with a colour/luminance most dissimilar to the attended items (see Figure 5.5). In Experiment 2, the attended items were at the centre of the luminance continuum used in Experiment 1, while the ignored items were at either end of the continuum. The cross was either the same luminance as the ignored items or at the opposite end of the continuum. Results showed that when the luminance of the unexpected object was the same as the ignored object, detection rates were much higher than when the unexpected object had a luminance at the opposite end of the continuum. Although this could suggest that selective ignoring contributes to inattentional blindness, the result could also be due to observers maintaining a luminance threshold, where ignored items fall on one side and attended items fall on the other (Most et al., 2001). In Experiment 3, the unexpected object was made more salient compared to the other objects by being red and having the other objects be circles or squares, rather than Ls and Ts. However, results showed that the cross was noticed less in this experiment than in Experiment 2.

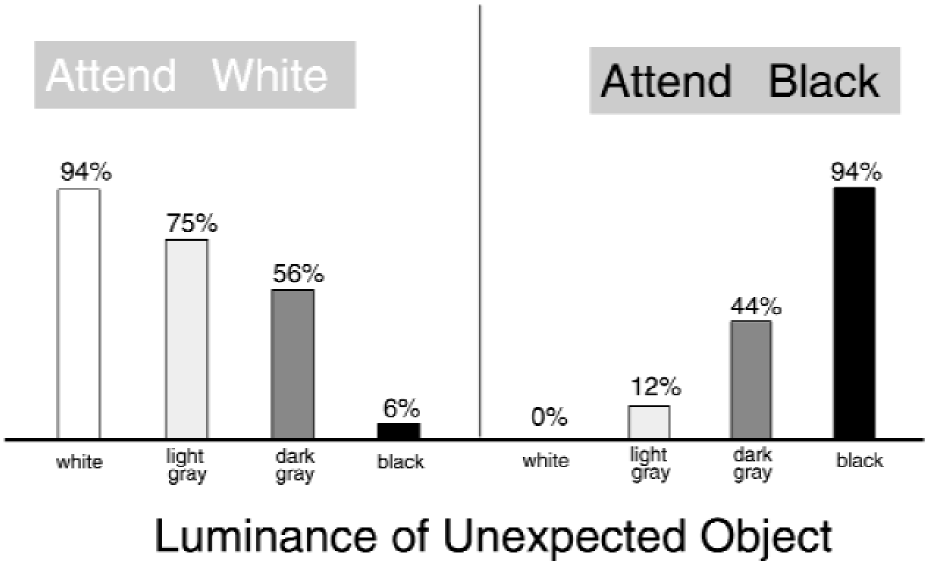


Figure 5.5: Results from Experiment 1 of the Most et al. (2001) study, showing results for different luminances of the unexpected object for conditions where participants attended to either the white or black objects.

5.3 Change Blindness

5.3.1 Changes During a Saccade

Inattentional blindness involves a failure to detect the presence of an unexpected stimulus, while *change blindness* involves a failure to detect a change to a stimulus. Change blindness received a great deal of interest in the late 1990s, after Grimes (1996) found that observers did not notice large changes made to photographs when these changes were made to coincide with the observers' saccades. Grimes (1996) instructed observers they were to study a photograph for an upcoming recognition test and that it would change periodically. The photograph changed during some saccades made by the observer, but observers did not notice most of the changes that occurred. This poor performance occurred despite the fact that changes were quite visually significant, such as the hats of two men being swapped. In another saccade-contingent change experiment, McConkie (1979) found that observers do not notice when words presented in alternating case[1] change as they make eye movements - the cases of the letters swapped each time

[1]Each letter in these words is the opposite case to the previous letter - ThIs Is An ExAmPlE.

the observer made an eye movement. More recent studies of change blindness do not make changes contingent on the observers' saccades, but instead interrupt the presentation of the stimulus when the change occurs.

5.3.2 The Flicker Paradigm

Rensink, O'Regan, and Clark (1997) made changes to photographs coincide with a blank screen, rather than with a saccade. The blank screen was used as a 'simulated saccade', as during saccades visual input is greatly suppressed (Burr, Morrone, & Ross, 1996). This paradigm is called the 'flicker' paradigm and is depicted in Figure 5.6. The flicker paradigm involves a picture (A) alternating with a blank field. A modified version of the picture (A′), in which an object in it or a portion of it has been changed, is presented after the blank field and the sequence (A, blank, A′) repeats until the observer indicates they have noticed the change. During the viewing time, observers are free to explore the scene with their eyes. Even when changes occur to significant portions of the visual scene, observers perform very badly on this task, taking up to several minutes to notice the change (Rensink et al., 1997).

5.3.3 Central and Marginal Interest - Semantic Determinants of Change Detection

In the Rensink et al. (1997) experiment, two characteristics of the change were manipulated - its type (colour, location or presence/absence) and its interest to the observer (central or marginal). The level of interest was defined by the results of a separate experiment in which observers described each scene - an object was of central interest if it was mentioned by three or more observers and was of marginal interest if it was mentioned by none. The average detection time was greater for changes made to objects of marginal interest than for changes made to objects of central interest. Location changes were significantly more detectable than colour and presence/absence changes when they were of marginal interest, but all three change types were equally detectable when they were of central interest.

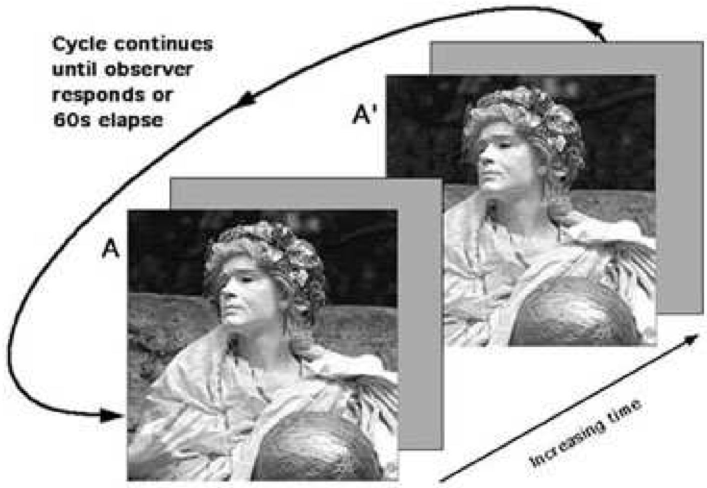


Figure 5.6: An example of a change blindness experiment run using the 'flicker' paradigm. From Rensink et al. (1997). The change occurs to the wall behind the statue, which decreased in height from A to A'.

The difference between central and marginal interest as defined above is probably related to both visual and semantic variation. Rensink et al. (1997) equated colour, luminance and size for central and marginal interest, in an attempt to control visual variation. However, given that only one change occurred per scene, central and marginal interest changes were always made in different scenes, meaning their relationship to the scene context would vary - this may affect a higher level of visual salience, or it may be a semantic difference. To overcome this limitation, Wallis and Bülthoff (2000) conducted a simulated driving study in which central and marginal interest changes were made in the same continuous (i.e., video-based) scene and were fairly similar in terms of low-level visual properties. The 'central interest' changes were relevant to the driving task, while the 'marginal interest' changes were not. Central interest changes were detected more readily than marginal interest changes.

Kelley, Chun, and Chua (2003) also controlled the central/marginal interest factor by using scenes in which two colour changes occurred simultaneously. These scenes were presented to participants who were required to respond as soon as they detected the change, and then box the change using the computer mouse. For a given scene, the change that was identified more often was labelled 'high' interest and the other was labelled 'low' interest. Both changes occurred to and from the same colour, and so should be relatively similar in terms of visual salience (although the spatial extent of the change could vary somewhat). In addition Kelley et al. (2003) presented half of the images to each observer inverted and half upright. Inversion of the images was predicted to disrupt global scene context, and so should decrease the preference for high vs. low interest changes. This was found to be the case, with high interest changes being detected 81% of the time in upright scenes and 69% of the time in inverted scenes.

However, Shore and Klein (2000) used a flicker paradigm and found that when the scenes are inverted, change detection RTs do not differ significantly from when the scenes are upright. Likewise, Yokosawa and Mitsumatsu (2003) found that disrupting the scene by rearranging sections, or removing sections (see Figure 5.7 for examples) did not impair change detection RTs. These results suggest that change detection does not depend on semantic factors, especially the global semantic context of the scene. The Rensink et al. (1997) and Kelley et al. (2003) studies suggest semantic content is important for the detection of change while the Shore and Klein (2000) and Yokosawa and Mitsumatsu (2003) studies suggest it might not be. However, in the Rensink et al. (1997) study, the control of visual salience was contaminated by changes in the scene context. This underscores the need for control of the visual information during the change.

5.3.4 The Mudsplat

In another type of change blindness experiment, visual input is not interrupted during the change but, instead, a large splat or group of dots is presented on the screen. In these 'mudsplat' experiments, the change may be visible, but the

Figure 5.7: Left: An example of a jumbled scene and Right: An example of a scene with portions deleted. From Yokosawa and Mitsumatsu (2003).

observer is unable to localise it due to the distraction of the mudsplat (O'Regan, Rensink, & Clark, 1999) (see Figure 5.8 for an example stimulus). The change blindness effect has also been shown when the changes are made gradually over time (e.g., when the change involves adding something to the scene, it can be made by 'fading-in' or increasing the contrast of the changing region gradually over time) (Simons, 2000). Therefore, the change blindness effect has been found when the change occurs during a saccade, a flicker, a mudsplat or when it is made gradually.

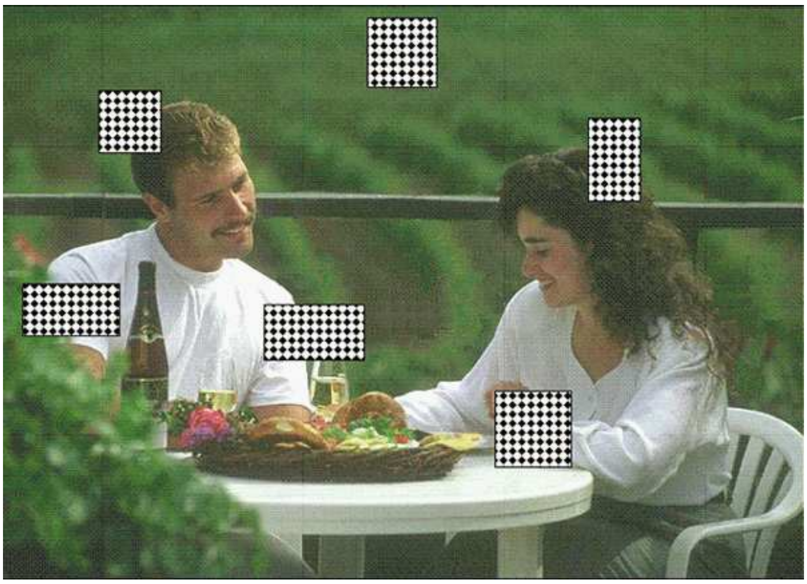


Figure 5.8: An example of a scene presented with mudsplashes. In this case, the change was the raising and lowering of the bar in the background. Each time the change occurred, the 6 checkerboard patterned boxes (mudsplashes) would appear. From O'Regan et al. (1999).

5.3.5 Changes to Multiple Objects

Rensink (2002) used the flicker paradigm to present observers with two objects, one or both of which could change depending on the experimental condition. Change detection performance was examined as a function of display duration, with possible change types being colour, size and orientation. For displays of less than 200ms, detection of pairs of changing items was always poor. For longer durations, performance improved for pair changes made to different features (e.g., colour and size) and also when both items were changing on orientation, but not for other types of change (i.e., colour and colour, size and size).

5.4 Motion and Change

All the different change blindness paradigms in some way prevent the motion signal normally associated with a change being localised by the observer. Motion can be defined as the temporal variation at a point in space of any measurable quantity (Adelson & Bergen, 1991). Therefore, motion occurs when something in a scene changes, because the change necessarily involves the variation over time of a measurable visual quantity such as colour and/or luminance. From this definition, motion can be considered as variation with reference to space. Change, by contrast, is variation with reference to structure - it involves the transformation of entire objects or parts of objects over time (Rensink, 2000c). However, because this variation necessarily involves the variation of low-level quantities such as colour and luminance, a change to a visual scene creates motion in that part of the scene. The motion signal created by a change is often called a *transient* because it occurs over a short period of time and is spatially restricted.

Transients will often orient the observer's attention to their location automatically (i.e., without the observer's volition) and this process is known as *attentional capture* (Rauschenberger, 2003). The literature on this phenomenon shows that while there is some evidence that salient objects, sudden onsets and some motion signals can automatically draw attention, it is difficult to rule out the involvement of top-down signals in the capture process (Rauschenberger, 2003). Regardless of how automatic capture of attention by a transient is, it is clear that the various change blindness paradigms all somehow interrupt or prevent the detection of the transient accompanying the change. The saccade-contingent changes and the flicker paradigm interrupt visual input during the change. The mudsplat paradigm creates a whole series of motion transients across the visual field and so no single one will capture attention. Because the transient signal can only occur if the change is made quickly, the gradual change paradigm does not create one.

It is clear that somehow obscuring the motion transient is necessary to ensure change blindness occurs. This does not, however, explain what causes change blindness from a psychological perspective - it says nothing about the internal processes that fail or are otherwise involved in the phenomenon. Giving the observer information about the change, such as what type of change will occur, or where the change will occur (i.e., a cue) can help elucidate the processes underlying the phenomenon.

5.5 Cuing the Change

As mentioned in Section 1.2, Aginsky and Tarr (2000) used the flicker paradigm to make changes to complex scenes and, before half of the trials, presented a cue word - either ‘colour’, ‘location’ or ‘appearance/disappearance’ that identified the type of change that was going to take place in that trial. Only colour changes were detected faster in the cued than in the uncued condition. The authors suggested that only colour changes were facilitated by cuing because the other changes are configural - that is, the other changes produce changes to the structure or configuration of the scene, while colour changes only change the surface of an object. They suggest that the configural properties of a scene are more readily encoded in memory than the surface properties, and so cuing makes the colour more salient and more easily transferred into memory.

Becker, Pashler, and Anstis (2000) conducted a flicker change blindness experiment with a circle of six letters in which the position of the one undergoing change was cued during the blank. Participants had to respond by first indicating if there was a change and, if they indicated there was, pointing to its location and saying what the pre-change number had been. This was done to test their ability to both locate and identify the change. Results showed that presenting a cue during the blank significantly improved participants' ability to locate and identify the change. However, in the next experiment, the cue was presented simultaneously with the post-change display. This failed to improve location and detection performance. The first result suggests that there is an internal representation (possibly iconic memory) of much of the scene, that remains present during the blank. The second result suggests this representation is overwritten by the post-change presentation.

The first study (Aginsky & Tarr, 2000) suggests that changes to surface and configural features are qualitatively different - the nature of a visual feature can affect detection of changes to it. The second study (Becker et al., 2000) involved the detection of changes to numbers, which could be considered a surface property, as it involved changes to the fine detail (identity) of numbers, without changing much spatial information. This study again showed that such changes are affected by cuing, but the cue this time was a location, rather than a type cue. The second study also suggested a cause for the change blindness - the overwriting of the representation of the initial stimulus presentation. Other causes for the phenomenon have also been suggested. However, it is important to note that, given the evidence from Aginsky and Tarr (2000), different features may be subject to different types and/or extents of change blindness - perhaps, then, failure to detect changes in them is caused by different internal processes.

5.6 Psychological Causes of Change Blindess

Simons (2000) listed the following as possible causes of the phenomenon:

- Overwriting of Representations - A′ overwrites A
- First Impressions Last (Initial Representations Persist) - A overwrites A′
- Nothing is Stored - a lasting perception/representation of either A or A′ isn't formed
- Nothing is Compared - A and A′ are stored separately and aren't compared
- Features are Combined across Representations - features from A and A′ are combined into a composite representation/perception

Similarly, Simons and Rensink (2005) suggest that change blindness experiments must eliminate the following possibilities before the common explanation for change blindness, that of 'sparse representations' can be validated:

- Detailed and complete representations exist, but decay or are replaced before change perception occurs
- Representations of the pre-change stimulus exist and persist, but are in pathways inaccessible to the mechanisms used for change perception
- Representations of the pre-change stimulus exist but are in a format that cannot be used for change perception
- Representations of the pre-change stimulus exist and are in an accessible pathway but the comparison operation is not applied

The 'overwriting' idea was addressed above in the discussion of the Rensink et al. (1997) and Becker et al. (2000) studies. The 'nothing is stored' idea simply means that no information from the stimulus is stored in any lasting fashion - as would be the case if the 'world as an external memory' idea of O'Regan (1992) has any validity. Evidence is lacking for this explanation, however. The 'nothing is compared' idea suggests that information from the pre- and post-change displays is stored independently and so is not compared - therefore, the change is not detected. In evidence of this explanation, Daniel J. Simons and Schnur (2002) found that observers were unaware of a change until after they had been asked about the previous state of the stimulus. In regards to the 'feature combination' idea - there is no evidence for the combination of features from the pre- and post-change scenes in the change blindness literature, but there is evidence of features from different objects being conjoined (i.e., illusory conjunctions - see Treisman & Gelade, 1980).

Regardless of the exact cause of change blindness, it is clear that the representations involved are somehow fragile. It could be that they decay quickly across time, are sparse in terms of the spatial information they contain, or both. Based on the fact that the representation of the current visual input is fragile (easily and quickly overwritten) and that attention allows information from this representation to be processed further (i.e., visual selection), two theories have emerged to explain the induced blindness phenomenon - Wolfe's *inattentional amnesia* hypothesis and Rensink's *coherence theory*.

5.7 Inattentional Amnesia

Wolfe (1999) presented an alternative hypothesis to account for the findings of studies on inattentional blindness and change blindness. This hypothesis is motivated, in part, by a point first raised by Sperling (1960) - that a failure to report does not necessarily mean a failure to see. Studies of inattentional blindness find that observers are unable to report the unexpected stimulus after it has been presented. Therefore, there are two possibilities: that the unexpected object was not seen (the inattentional blindness hypothesis) or that it was seen and quickly forgotten. Wolfe (1999) adopts the later hypothesis which he calls the *inattentional amnesia* hypothesis. The hypothesis is based on two main ideas:

- The current visual representation includes information from everything across the visual field but this representation has no memory - changes in the visual field instantly erase information that was there before

- Attention enables the visual representation to make contact with other mental processes, such as object recognition and memory

It follows from these ideas that unattended stimuli may be seen, but not remembered, because only the locus of attention can enable information from the visual representation to be transferred into memory. This hypothesis also applies to change blindness, because it suggests the failure/blindness here is not one of perception (i.e., that the change is not seen), but of memory (i.e., that the change is seen but not remembered).

5.8 Coherence Theory

Rensink (2000b) constructed a theory based on his idea that focussed attention is needed to see change, which he called *coherence theory.* The theory proposes a visuo-attentional architecture consisting of an attentional system and a non-attentional system. The attentional system serves the purpose of generating coherent objects from an assortment of 'proto-objects' across the visual field. These proto-objects are generated in the non-attentional system through perceptual processes such as grouping, image segmentation and surface completion[2]. They are volatile structures, meaning they are generated quickly and can disappear just as quickly - they are unstable. Their generation is supposed to be automatic and proceed quickly across the visual field. It is known that grouping and surface

[2]These are *perceptual processes* in the sense defined previously - processes that deal with the physical visual qualities of objects and are only mildly, if at all, influenced by attention or any other top-down signals. See Moore and Egeth (1997) for more detail.

generation in early vision are very quick, automatic processes (Moore & Egeth, 1997), and are therefore candidates for the formation of proto-objects. Subsequent to the formation of proto-objects, the attentional system acts to select a subset of them in a 'coherence field' (cf. focussed attention in Chapter 3, Section 3.6.1) to create a complete object representation, while the visual space outside the coherence field is represented only in schematic form.

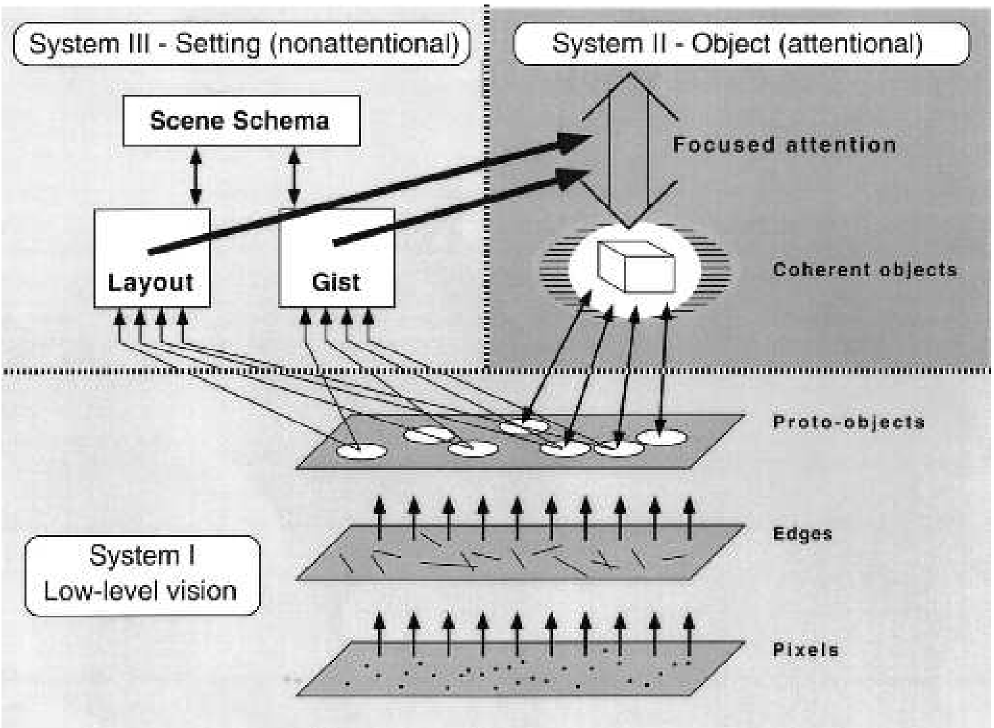


Figure 5.9: A schematic of Rensink's coherence theory of visual attention. From Rensink (2000b).

Coherence theory, like Feature Integration Theory (see Chapter 3) proposes that elementary features (proto-objects) are integrated by attention into visual objects. However, neither the theory nor the studies on change blindness discussed above can quantify how different changes in different features contribute to the overall perception of change. To attempt such a question, the stimulus and the changes made to it must be more controllable than they are in the context of complex scenes typically used in the change blindness studies. In Chapter 3, the visual search paradigm was discussed - this paradigm offers a way to control factors such as the number of objects present and the values those objects have on various featural dimensions. Although the Becker et al. (2000) study was a visual search task, it changed the elements on an abstract dimension - the identity of the numbers, rather than on a lower-level dimension such as colour or orientation.

5.9 A Search Paradigm for Change Detection

Rensink (2000d) combined the visual search and flicker paradigms to create a 'visual search-for-change' task, in which, on half of the trials, one of the elements changes either its orientation or its contrast. There were two values on each of these dimensions - orientation could be 0° or 90° and contrast could be black or white. When a change occurred, it was made from one of these values to the other. The two values for orientation and contrast were distributed roughly equally amongst the elements. A main aim of Rensink's study was to determine whether visual search for a change is similar to search of static patterns for a target defined by a difference to the distractors. To investigate this, the relationship between reaction time (i.e., time to respond correctly if target is present or absent) and set size was investigated for different target types and two different on-times (the amount of time the elements are presented for before and after the blank screen) - 80 msec and 800 msec. This relationship was linear for both types of change, and the slopes for target absent trials were about twice those of target present trials in each case, suggesting the search process was self-terminating (see Chapter 3, Section 3.3).

Search slopes for orientation were different for the two on-times, and so another experiment was performed to examine performance at several different on-times. It was found that search slopes only varied with on-times for orientation changes with on-times of 640 msec and 800 msec. Therefore, search slopes are relatively constant for changes in on-time. Rensink (2000d) suggests this means search speed is governed by intrinsic processing constraints rather than factors that would depend on display time, such as stimulus quality (e.g., how well stimuli can be resolved) or memory limitations. If this interpretation is correct, it is likely that search slopes for this task are related to an overall processing time, which Rensink (2000d) suggests would consist of: the time needed to load information in vSTM, hold it across a temporal gap, compare it, and - if necessary - unload vSTM and shift processing to the next candidate item. In Experiment 3, the relationship between the search slope and the 'hold' of objects across the temporal gap was investigated. It was hypothesised that the number of items held are compared with an item in memory once every alternation (on-time + blank/off-time), giving the following formula:

$$hold = (on\ time + off\ time)\ /\ search\ slope \tag{5.1}$$

where *hold* represents the average number of items held across each temporal gap. Rensink then defines capacity as the asymptotic hold that exists with increasing on-time. Data from Experiment 3 is shown in Figure 5.10, and shows that orientation reaches an asymptote at about 5 items but polarity does not reach one. It is possible, therefore, that polarity is a feature that can be grouped across elements, or that hold must be considered in the context of individual features rather than objects. This matter can only be resolved by further study using this paradigm.

It is clear from Rensink's experiments that a more controlled paradigm can better delineate the nature of change blindness and its relationship to visual search and visual memory. Furthermore, it shows a way of estimating the capacity of the attentional 'hold' for a 'visual search for change experiment'. H. Pashler (1988) also formulated a way of estimating capacity (in terms of the number of items that can be compared) for change detection when accuracy is the dependent measure (see Equation 5.2). This equation has been used in many change detection studies to estimate the capacity of vSTM. k represents capacity, n represents set size and g is the probability of guessing that a change occured (i.e, the false alarm rate). The

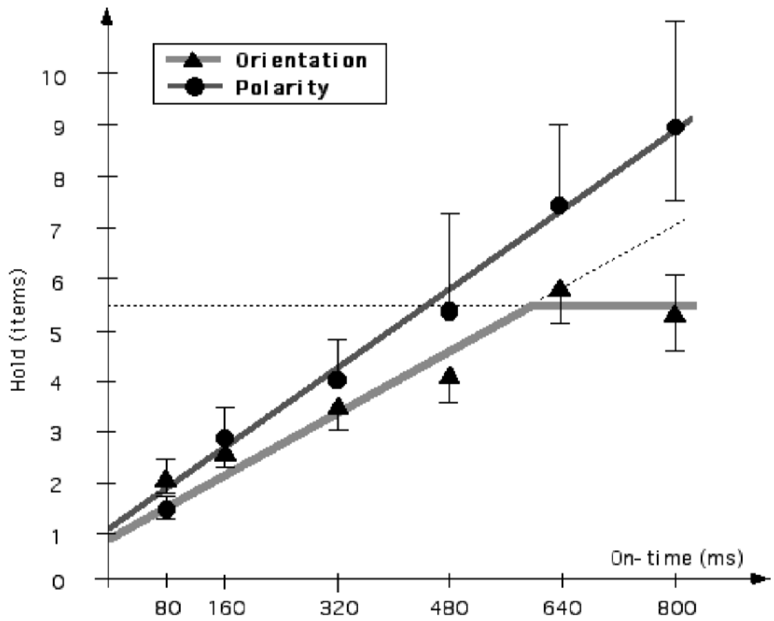


Figure 5.10: Data from Experiment 3 of the Rensink (2000a) study, showing hold as a function of on-time. Hold reaches an asymptote at 5 items for orientation, but does not reach an asymptote for polarity.

model assumes no partial information (e.g., parts of objects) and also attributes all errors to a maintenance, rather than a comparison process (see Section 5.6).

$$H = \frac{k}{n} + \frac{(n-k)}{n * g} \tag{5.2}$$

5.9.1 The Experiments of Brown and Orbach (1998) and Brown et al. (2000)

Brown and Orbach (1998) used a 2AFC psychophysical paradigm to more precisely quantify change detection. They presented participants with two sets of gabors, in succession, around the fixation point. On half of the trials, one of the gabors changed its contrast from the first presentation to the second. Participants were presented with single, three and five element patterns and were required to identify which interval contained the element that had increased contrast. Results were reported as contrast increment thresholds as a function of base contrast (of the changing gabor) and set size. Figure 5.12 shows the results of this study -

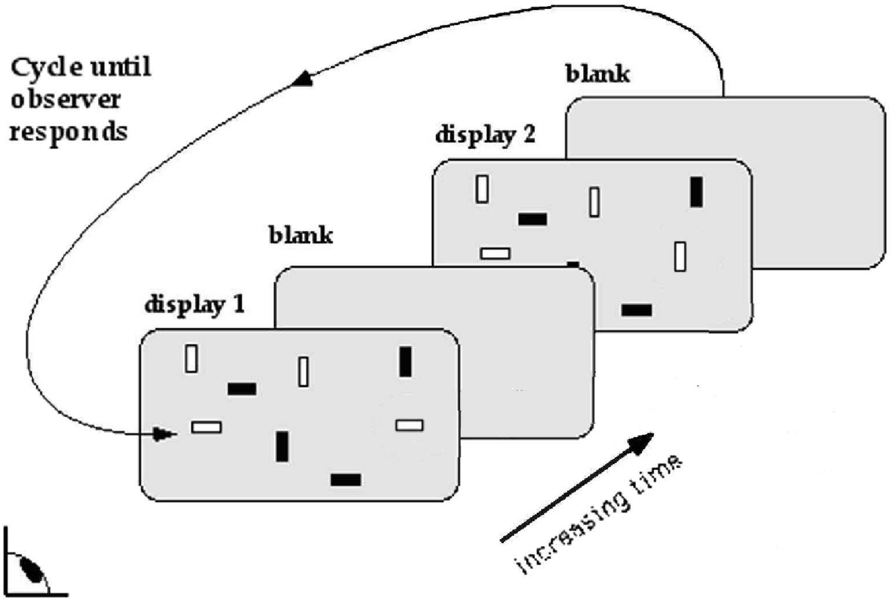


Figure 5.11: A schematic of Rensink's visual search-for-change task. Adapted from Rensink (2000d).

thresholds increased with set size. Figure 5.12 also shows the set size effects of the mixed (heterogenous distractors) and uniform (homogenous distractors) conditions, and compares them with the set size effect obtained in the visual search contrast increment experiment of Palmer (1995). Although the uniform condition shows a similar log-log slope to that obtained from Palmer's experiment, the mixed condition shows a much larger effect. Brown and Orbach (1998) argued that these results could not be successfully explained by either a limited capacity model nor a decision-noise limited spatial allocation model. Although the mean log-log slope obtained for the mixed task (0.63) is similar to the slope predicted by a limited capacity model (0.75), the limited capacity model fails to explain the results of the uniform condition. Consequently, Brown and Orbach (1998) suggest that change blindness is caused by factors outside of 'early' attention.

In a subsequent study, Brown et al. (2000) asked whether the effect would be observed when the patterns are presented simultaneously, rather than successively (see Figure 5.13). Therefore, this study included a simultaneous presentation condition, in which one pattern was presented to the left of fixation and one

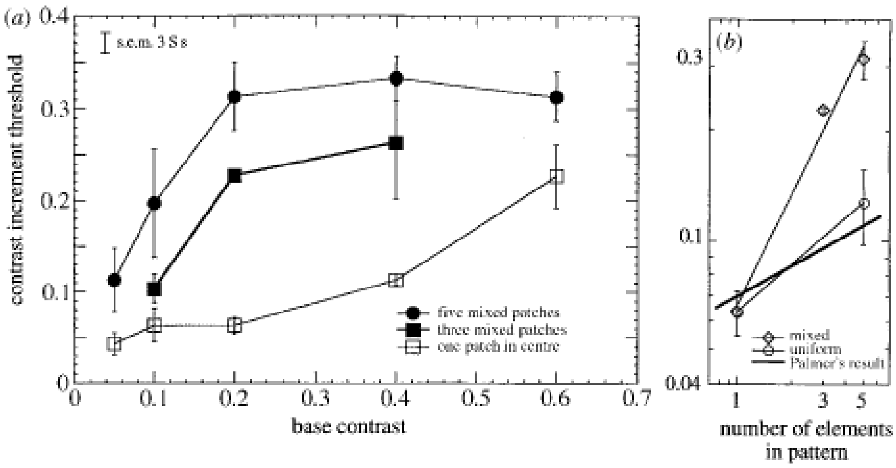


Figure 5.12: a) The results of the Brown and Orbach (1998) study. b) The log-log set size vs. contrast increment slopes of the mixed and uniform conditions and that obtained from Palmer (1995)

was presented to the right and a successive presentation condition was run for comparison. The gabors in the successive presentation condition were presented in the same positions as in the simultaneous presentation condition (i.e., one to the left and one to the right of fixation), to make the comparison between conditions as valid as possible. Each set was presented in a different vertical position so observers could not detect the difference by a simple break in symmetry, using low-level mechanisms. There was also a cued condition in which the position of the target was indicated by a white blob amongst black blobs (indicating distractors) in a miniature representation of the set of gabors presented at fixation. The presentation of each set of gabors in the successive condition lasted 250 msec, as did presentation of both sets in the simultaneous condition. A Wilcoxin matched-pairs test was made for each participant at each base contrast and it was found that there was no significant difference between simultaneous and sequential trials, when they were either cued or uncued. However, cued and uncued trials were significantly different from each other for both simultaneous and successive presentations.

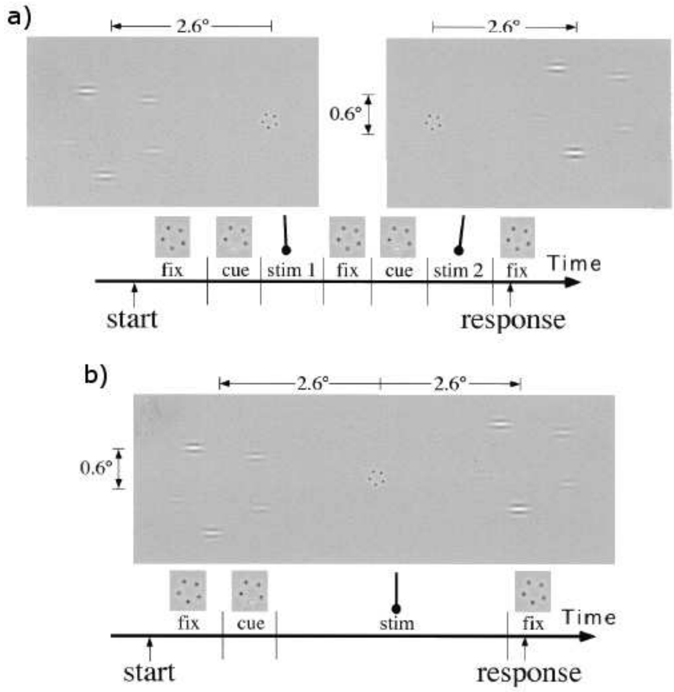


Figure 5.13: The a) successive and b) simultaneous presentation conditions of the Brown et al. (2000) study.

5.9.2 Summary of Search-for-Change Studies

The results of Becker et al. (2000), Rensink (2000d), Brown and Orbach (1998) and Brown et al. (2000) demonstrate that change blindness is not restricted to complex scenes. The lack of a difference between the simultaneous and successive presentation conditions in the Brown et al. (2000) study could demonstrate that change blindness is due to a failure of comparison, rather than one of misperception. Although the Becker et al. (2000) experiment suggested that change blindness is due to overwriting of the iconic representation, these two results can be reconciled by the idea that the overwriting means the pre- and post-change displays cannot be compared, as information from the former is not retained. Also, the overwriting could not occur in the Brown et al. (2000) study, because the pre- and post-change displays are in different locations (assuming the relevant representations are retinotopic). The Rensink (2000d) study partly examined the comparison process by looking at the capacity of vSTM for holding elements across the blank interval - the same question was looked at in the Luck and Vogel (1997) study discussed in Chapter 4, using different methods.

5.9.3 Signal Detection Models of Change Blindness

Wilken and Ma (2004) conducted a study using a signal-detection approach in which the capacity of vSTM was examined separately for colour, spatial frequency and orientation. The paradigm used was a search paradigm similar to that used by Brown and Orbach (1998) and Brown et al. (2000). Figure 5.14 shows a schematic of an example trial in which a colour change occurs. For orientation and spatial frequency changes, a ring of gabors was used instead of a ring of coloured squares.

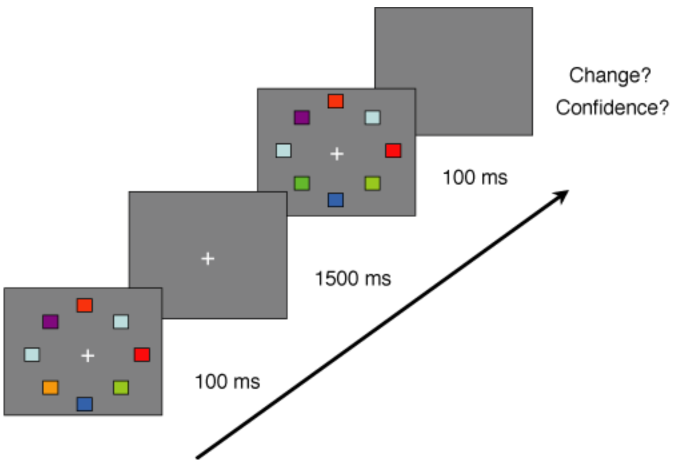


Figure 5.14: A schematic of a colour change trial in the Wilken and Ma (2004) study.

The Wilken and Ma (2004) study, like the experiments in Palmer et al. (2000), used signal detection theory to analyse results. Signal detection theory is a low-threshold theory, meaning that, in contrast to high-threshold theory, it allows for the possibility of false alarms (i.e., indicating there was a target when there wasn't) due to the threshold for detection being *low* enough to be reached by noise alone (i.e., when no signal is present). More traditional theories dealing with visual search and vSTM are high-threshold theories. Such theories generally posit that the patterns of performance observed are due to a limited capacity store that can contain a fixed number of discrete items or objects (e.g., see Equations 5.2 and 5.1 and related discussion). Low-threshold theories also allow for vSTM being a limited capacity store, but simply suggest that the representations of the item/objects will be noisy and not discretely broken up into a number of pre-allocated 'slots'. Instead of predicted performance in terms of the number of items, then, low-threshold theories make predictions in terms of measures related to sensitivity such as hit rates and false alarm rates. In the Wilken and Ma (2004) study, data were fit to a high-threshold model (provided by Equation 5.2) and also to two low-threshold signal detection models - the maximum of absolute differences (MAD) model and the sum of absolute differences (SAD) model. The key differences between the high- and low-threshold models are that the low-threshold models assume a noisy representation and also give a rule for how the decision stage integrates across multiple informational channels (i.e., with a sum in SAD and a maximum rule in MAD). Data for the Wilken and Ma (2004) was better fit by the low-threshold models than the high-threshold model, and of the low-threshold models was best fit by the MAD model.

5.10 Conclusion

Although the signal and noise may be defined at a higher level in a search or change blindness experiment, it does not necessarily mean that the threshold for detection of the signal is high, as compared to the noise. Therefore, it may be more useful to view the search and change blindness tasks in the context of low-threshold theories like signal detection theory. This may help resolve some of the outstanding questions in the change blindness and search literature. For instance, what are the relative contributions of overwriting and failures of comparison to the change blindness effect? Also, how is the capacity of vSTM to hold and compare objects affected by different features changing, and different magnitudes of change? At the very least, the studies dealt with in the last section of this chapter demonstrate how a more controlled paradigm can be used to better analyse the processes underlying change blindness. It is for this reason that the current thesis used a paradigm derived from these studies for all of the experiments conducted.

Chapter 6

Rationale for the Experiments and Overall Predictions

6.1 Rationale

All of the experiments of this thesis are visual search experiments, in which the observer is presented with a number of spatially separate and distinct elements on each trial and is required to detect a target or discriminate a property of a target from other properties of the target or properties of non-target elements (distractors). Having participants search through a series of elements rather than a complex scene allows greater control over the stimulus, and this can assist in the theoretical interpretation of the data.

The main reason for using an experimental paradigm that gives greater control is that it enables comparison of changes in different stimulus dimensions in a meaningful way. As was discussed in the Introduction, most change blindness research has used complex, naturalistic scenes as stimuli, and the changes that were being made to these scenes usually involved transformations or the addition or deletion of complete objects. It is unclear from such studies what properties of the change are important in its detection (or lack thereof) and how different properties contribute to the overall detection and identification of change. The visual search paradigm is usually used to answer questions regarding how different features and properties contribute to the detection of a target or event amongst distractors.

Therefore, the experiments in this thesis applied visual search methodology to the problem of change blindness so that more specific questions could be answered regarding the relative contributions of different features to the change blindness process and also regarding whether it is object-based or feature-based encoding that underlies the phenomenon, or perhaps a combination of both.

All of the experiments in this thesis use elements that are presented in a ring around a central fixation cross, and these elements do not change their positions during the trial. All of the experiments control both the eccentricity of stimuli relative to fixation, and the eye movements of observers, by requiring observers to fixate and by presenting stimuli for brief times so that multiple eye movements cannot be made during any given presentation of a set of elements. Using this method constrains variability of the data and allows us to apply psychophysical techniques to analyse it (see Palmer et al., 2000). This also allows interpretation of the data in a more fine-grained manner, dealing with processes at a level relatively independent of the high-level visual and semantic context of the scene.

Another method used in the current experiments to constrain variability is the use of accuracy, rather than reaction time, as a dependent measure. This means that the stimulus is presented for a fixed amount of time in each trial. Furthermore, using accuracy makes it possible to make inferences regarding the capacity limitations of the processes involved in search, whereas this is not necessarily the case when reaction time is used as the dependent measure (see Section 3.3.5.2).

Using psychophysical methods to examine a phenomenon like change blindness allows us to apply tried and tested tools to a phenomenon that still raises many unanswered questions. By using controlled presentation and analysis paradigms, we can more clearly answer such questions and offer a more precise statement of unresolved issues. However, the studies of change blindness using complete visual scenes are still of major importance in guiding the design of such experiments and help in understanding the role of visual context in the phenomenon. Therefore, the relationship between these two approaches to change blindness will be examined in the Discussion.

6.2 Predictions

Given that these experiments are all similar in terms of the stimuli used and the responses required of observers, general predictions can be made concerning the outcomes of the experiments. Firstly, it was predicted that the results would support a signal detection approach to change detection, in preference to high-threshold models. This was expected to come in the form of the sensitivity vs. set size data conforming to the sample-size model (see Section 3.3.5.1). Furthermore, it was expected that log-log slopes generated from the accuracy vs. set-size data would be better described by low threshold theories than by high threshold theories (see Palmer et al. (2000) and Section 3.4.3.2). It was also expected that performance on the change detection tasks dealing with multiple features and objects changing simultaneously would support a model of change detection where it is the number of features changing that dictates performance (see Wheeler and Treisman (2002), Wilken and Ma (2004)), rather than the number of objects.

Chapter 7

General Method

In a typical visual search experiment the target is different from all of the distractors on some featural dimension such as colour, size, orientation or spatial frequency. In all but one of the experiments in the current study, however, targets were elements that underwent a *change* along some featural dimension across a blank interval (i.e., before the blank interval, they were unchanged and after it, they were changed). This paradigm has been called 'visual search for change', as it involves searching for a target that is defined by a change (Rensink, 2000a).

7.1 General Procedure

Participants were recruited from either the University of Queensland Human Movements Department, the University of Queensland Psychology Department, or an online employment website for students of the University of Queensland. When a participant came for an experiment, they first had to take a basic colour vision and acuity test, which was run using the Bausch & Lomb Vision Tester. The colour vision test used four pseudoisochromatic plates that were corrected for instrument illumination. The acuity test used a series of non-letter stimuli (numbered squares containing diamonds with shaded patches inside) arranged in rows where each square was progressively smaller than the last. The observer was required to identify the location of a checkerboard within each numbered square. These tests were administered using a scoring card and participants were required to fulfill recommended criteria in order to continue to the experimental phase. If participants passed this initial testing, they were seated in front of the computer on which the experiments were run and were guided through a training session. Training sessions were run until the observer indicated they were comfortable with the task and were performing it properly.

7.2 Equipment and Stimuli

The experiments were run on an SGI Onyx 300 machine using custom software. A Silicon Graphics (Sony Trinitron) CRT monitor with a display size of 43 cm x 29 cm and resolution of 1280 x 1024 (fully anti-aliased) was used, with participants sitting with their eyes 100 cm from the surface of the display. The display therefore subtended 24.26 x 16.50 degrees of visual angle.

In all the experiments, participants were presented with a ring of gabors (a gabor function is shown in two- and three-dimensions in Figure 7.2) and required to perform a discrimination, detection and/or an identification task[1]. The stimulus subtended a radius of 5.0 degrees, meaning the centre of each gabor was 5.0 degrees from the centre of the fixation cross, and 5.0 degrees from the centre of either of the two adjacent gabors. Gabors were constructed using Equation 7.1 as a basis for each of three different colour channels (see Appendix 14, Section 14.1 for the source code used to construct the gabors).

$$sin(cycles \times 2\pi \times x/size + \phi)^{(size/2-x)(x-size/2)+(size/2-y)(size/2-32)} \tag{7.1}$$

where *cycles* is the total number of bars in the gabor, *size* is the texture size (constant at 64 texels) and ϕ is the phase offset of the gabor in radians.

All experiments involved variation along one or more featural dimensions in each trial. The program controlling stimulus presentation explicitly coded size,

[1]A discrimination involves comparison between two or more stimuli (elements), while detection and identification each involve a single one, but detection requires only 'present' or 'absent' response, while identification requires a response containing semantic information.

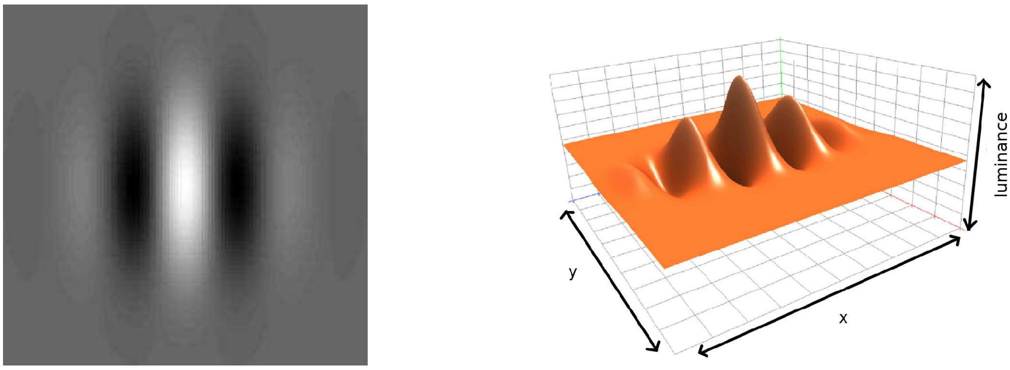


Figure 7.1: On the left is a gabor function in two dimensions (as it appeared on the screen) and on the right is the same function with an added dimension of relative brightness.

orientation, colour on the red-green axis, speed of motion, direction of motion and spatial frequency. Speed of motion was temporal frequency. Temporal frequency is the number of cycles that pass a given point per second. It is related to phase by the equation: t.f. $= \phi/t$. Temporal frequency is the speed at which sinusoids within the gabor moved (i.e., the gabors themselves remained stationary). In real-world visual stimuli, many featural dimensions co-vary (e.g., colour and luminance). To ensure that covariation along featural dimensions could be controlled, the current study employed the following measures:

- To control covariation of colour and luminance, the peak colour values for the gabors were made isoluminant (lightness of 22) - see Section 7.2.1, below.

- Given that, for any given speed, most motion information is recoverable when the object is moving in a direction orthogonal to its orientation (Clifford, 2002), the movement of the gabors was constrained so that the gabors could only move in the two directions orthogonal to their orientation.

- The gabors were always defined with a gaussian window of 1.8 degrees or more, and at least two bars were visible. This was to ensure sufficient motion information was available in all different forms in which the gabor appeared (i.e., to ensure that for changes in spatial frequency and/or size, perceived speed still varied with actual speed). Also, in experiments where speed judgements were critical, the upper limit on the spatial frequency of gabors was constrained so that the bars were always thick enough (i.e., ϕ was always large enough) to yield sufficient motion information (when spatial frequency is too high, motion appears as a directionless flickering). These constraints on spatial frequency and size were made using judgements and adjustments of several psychophysical observers.

7.2.1 Determining Colour Points

Because the colours used had to be isoluminant, there was no on-line modification of colour values in any experiment - all colour values were preset. To do this, the red and green values of the RGB computer colour space were varied so that a series of points were created that were progressively less red and more green, but all of the same lightness value (as read by a Minolta CL-100 colorimeter operating in CIE 1976 L*a*b* colour space, where L is lightness). The use of this procedure means that the space between consecutive colour points is not uniform on a metric scale - the placement of consecutive points was instead made on the basis of colour appearance. In other words, points were created so that each subsequent point appeared (to the experimenter) to be as distinct from the previous as other pairs on the scale and all points were made isoluminant. Each colour point was measured in the CIE 1976 L*a*b* colour space[2] where L is lightness, and a and b are colour co-ordinates on each of the two colour-opponent dimensions (a is red-green and b is blue-yellow, see Figure 7.2). These co-ordinates were converted as shown in Equation 7.4 to give values in the L*C*h colour space (C and h represent chromaticity and hue) so that each colour point could be described as a single point - its h value.

[2]This colour space is also referred to as ‘L*u*v*’ and ‘CIELAB’. It is derived from the CIE 1931 XYZ colour space and was created to be more perceptually relevant and uniform than the 1931 colour space.

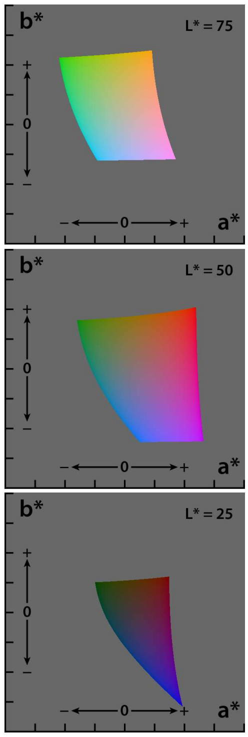


Figure 7.2: The CIE 1976 L*a*b* colour space, for three different lightness (L) values, showing the colours that would be perceivable on a standard monitor. Both axes go from -128 to +128. a is the Red-Green colour axis and b is the Blue-Yellow axis.

$$L(lightness) = L \tag{7.2}$$

$$C(chromaticity) = \sqrt{a^2 + b^2} \tag{7.3}$$

$$h(hue) = tan^{-1}(b/a) \tag{7.4}$$

Chapter 8

Changes to Gabor Stimuli - Display Set-Size and Relevant Set-Size

8.1 Experiment 1 - Changes to Stationary Gabor Stimuli

8.1.1 Questions and Hypotheses

The main question motivating this experiment was:

- How do different change types compare in terms of their relative detectabilities?

A secondary question was:

- Can set-size effects in this paradigm be predicted by the sample size model, as used by Palmer (1990)?

The experiment was designed to compare detection performance for changes in speed, size, orientation and colour. The magnitude of changes was constrained to a range of values that produced rougly equal levels of performance across change types in pilot testing. It was predicted that:

1. Changes to different stimulus dimensions (speed, size, orientation and colour) would produce similar set-size effects, in agreement with Palmer's (1994) finding that visual search for increments in basic features such as disk colour,

disk size, ellipse orientation and speed discrimination produce similar set size effects and also Wilken's (2004) finding of similar set-size effects for changes in colour, orientation and spatial frequency.

2. Performance would increase with change magnitude, in accordance with the findings of Brown and Orbach (1998) for changes in the contrast of gabors.

3. Set-size effects would follow the relationship:

$$d' \propto \frac{1}{\sqrt{N}} \tag{8.1}$$

as shown by Palmer in his studies of visual search for simple features. This relationship is reflective of a simple decision noise model and can be contrasted with a high threshold model (see Equation 5.2), which predicts the relationship:

$$d' \propto \frac{1}{N} \tag{8.2}$$

8.1.2 Method

8.1.2.1 Participants

Four participants took part in this experiment. All were female undergraduate students recruited through a website at the University of Queensland. Their mean age was 20.75 (range 18-26). Each was paid $10 for their participation.

8.1.2.2 Display

On each trial, the participant was presented with a number of coloured gabors. The gabors differed in colour, size and the orientation of their textures. Furthermore, each gabor was moving in a random direction (left or right) normal to its orientation and the gabor textures could move at different speeds. On half of the trials, one of the gabors could change its colour, size, speed or the orientation of its texture. This change co-incided with a blanking of the screen.

8.1.2.3 Procedure and Design

The progression of a trial in this experiment can be schematised as:

$$fixation \rightarrow A \rightarrow B \rightarrow A' \rightarrow B \rightarrow A \rightarrow B \rightarrow A'$$

where A is the pre-change presentation, A′ is the post-change presentation and B is a blank screen. On trials in which no change occurred, the elements in A′ are the same as the elements in A. Fixation lasted 1500 msec, pre- and post-change presentations lasted 1500 msec. Blank screens were presented for 120 msec. Figure 8.1 shows an example trial.

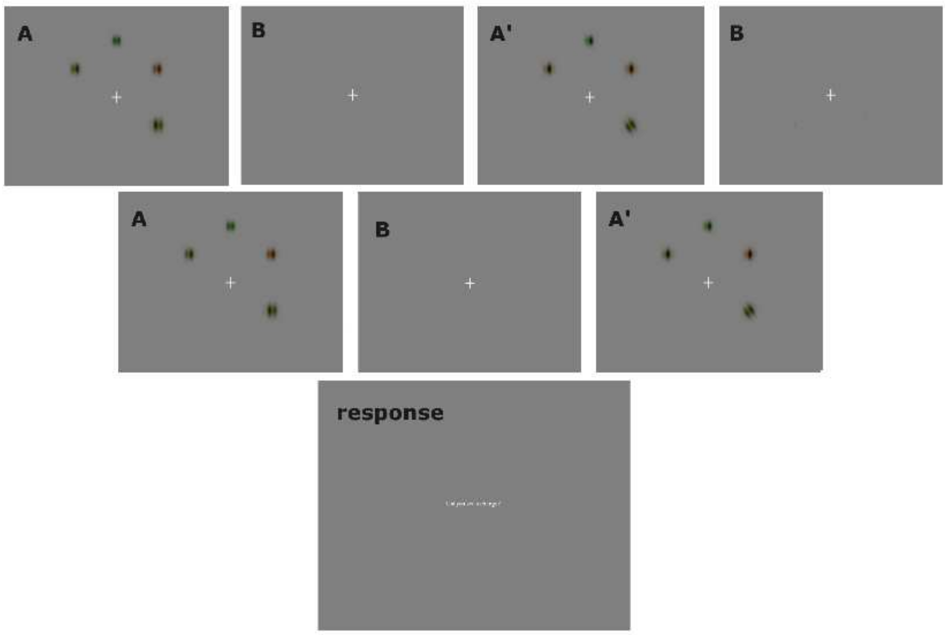


Figure 8.1: An example trial with an orientation change and a set size of four. The changing element is below and to the right of the fixation cross.

Independent variables in this experiment were: the presence of change (present or absent), the type of change (colour, speed, size, orientation), the set-size (1, 2, 4 or 6), and the magnitude of change (2-5 steps, where 1 step is equivalent to moving from one of the stimulus values to the neighbouring greater value).

Therefore, there were 64 unique trials in which a change occurred and 64 trials in which no change occured. The experiment was run in two blocks, each containing 128 trials (64 change and 64 no-change) run in a random order. At the end of each block, the participant could take a break. A message on the screen prompted them to press the enter key to continue to the next block.

8.1.2.4 Stimulus Configuration

In each trial, elements adopted a value for colour, orientation, size and speed selected randomly from the points listed below. The starting points of changing elements was constrained so that they could only change from one of the points listed to another (within a dimension).

- Color (hue): 0.98, 1.01, 1.04, 1.08, 1.11, 1.21, 1.24, 1.27, 1.30, 1.32
- Orientation (degrees): 0, 15, 30, 45, 60, 75, 90, 105, 120, 135, 150
- Size (degrees): 1.8, 2.07, 2.38, 2.74, 3.15, 3.62, 4.16, 4.78, 5.50
- Speed (degrees per second): 0.3, 0.47, 0.72, 1.12, 1.73, 2.68, 4.16, 6.45, 10.00

The luminance of elements was constant throughout (see Section 7.2.1). This was done so that variations in colour were independent of variations of luminance, in keeping with the psychophysical principle of varying one stimulus dimension independently of others (see Palmer, 1994). The values for speed and size lie on a logarithmic scale, and follow Weber's law such that the difference between two neighbouring values is proportional to the magnitude of the lower value (see Equations 8.3 to 8.5). This was done because it was found in pilot testing that the discriminability of successive increments in change magnitude remained the same only if the size of these increments followed Weber's law (for speed and size changes). Spatial frequency was fixed at 1.5 and gabors could drift left or right, relative to their orientation and this direction was randomized.

$$\Delta L / L = k \tag{8.3}$$

$$\therefore \Delta L = kL \tag{8.4}$$

$$\therefore \Delta L \propto L \tag{8.5}$$

8.1.3 Results

8.1.3.1 Hit Rate versus Proportion Correct

There are two obvious dependent variables that can be used to analyse this data: *hit rate* and *proportion correct*. Hit rate is the proportion of trials in which a change is correctly identified (i.e., change trials in which the participant indicates there was a change). Proportion correct is the hit rate added to the rate of correct rejections. A correct rejection is when a participant correctly indicates on a no change trial that there was no change. In addition, a False Alarm is when a participant indicates there was a change on a no change trial. Equations 8.6, 8.7 and 8.8 show how hit rate, proportion correct and the number of correct rejections are calculated, respectively.

$$Hit\ Rate = Hits\ /\ Trials \tag{8.6}$$

$$Proportion\ Correct = (Correct\ Rejections + Hits)\ /\ Trials \tag{8.7}$$

$$Correct\ Rejections = (Trials\ /\ 2) - False\ Alarms \tag{8.8}$$

Trials in which a correct rejection occurs are not unique in the way that hit trials are - they cannot be uniquely attributed to a category or 'cell' defined by a combination of IVs (change type, magnitude of change) as there is no change occurring in a correct rejection trial. Therefore, any difference in performance when using Proportion Correct as a DV rather than Hit Rate will be random. This is because the difference is due only to the correct rejection trials included in the analysis. These trials cannot be included based on the combination of conditions occurring in that trial (as the trials are exactly the same regardless of what condition the program/experiment may attribute them to), but on a purely random basis instead. Therefore, hit rate is the more meaningful DV to use for the current experiment.

8.1.3.2 Data

Based on the rationale outlined above, hit rate was used as the dependent variable for this experiment. Figure 8.2 shows how hit rate varies with set size for each of the four participants, and also gives the mean hit rate of the four participants. Figure 8.4 shows how the mean hit rate varies with set size for the four different change types. Table 8.1 gives the best fitting slope (from robust regression) and the corresponding error for each change type for set size vs. hit rate.

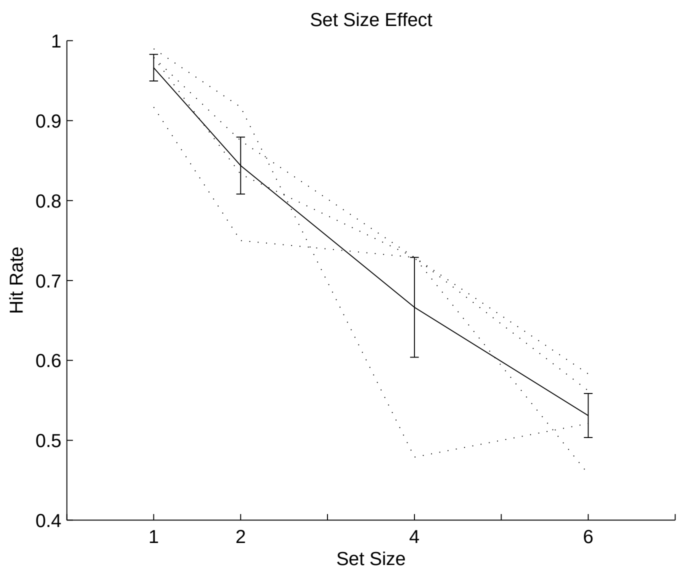


Figure 8.2: How hit rate varies with set size. Data for each of the four participants is shown as a dotted line and the mean is shown as the dark line, with standard errors included.

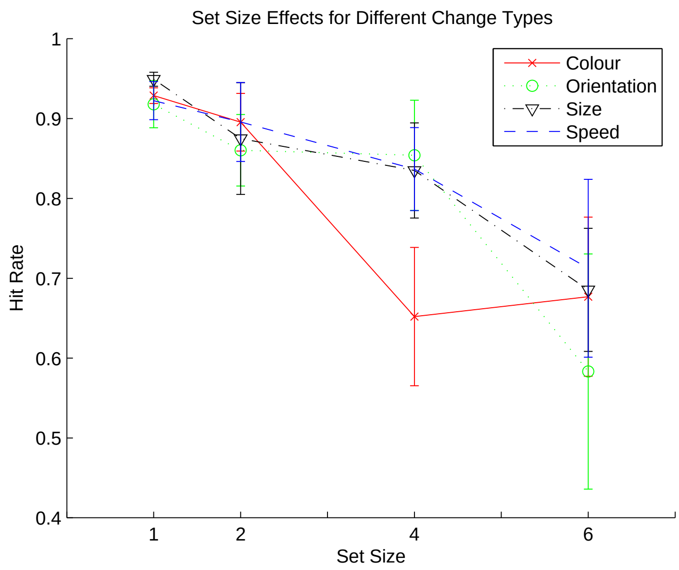


Figure 8.3: How hit rate varies with set size for the different change types, averaged across participants.

Table 8.1: Experiment 1: Set-Size Effect Slopes and Error

Feature Dimension	Slope	Error
Colour	-0.0584	0.0219
Orientation	-0.0605	0.0222
Size	-0.0488	0.0096
Speed	-0.0412	0.0064

It is clear from these results that performance is relatively similar for all change types across set size, the exception being colour changes at a set size of four, which have a significantly lower hit rate than the others. Orientation and colour have the largest set size effects (largest rate of performance decrement with increases in set size), and speed and size have lower set size effects that are similar to one another. However, the size and speed set size effects are within the error range of the other effects. Therefore, all four set size effects are relatively similar. Figure 8.4 shows how hit rate varies with change magnitude for the four different change types. All change types show an increasing hit rate with change magnitude for magnitudes of 1 and 2. However, hit rate for orientation changes decreases with magnitudes of 3 and 4, while the hit rate for the other change types continues increasing.

The data also show that performance at set sizes of one for each change type is close to ceiling, reflecting the fact that these changes are supra-threshold and easy to detect in the presence of no distracting information (which is the case for a set size of 1).

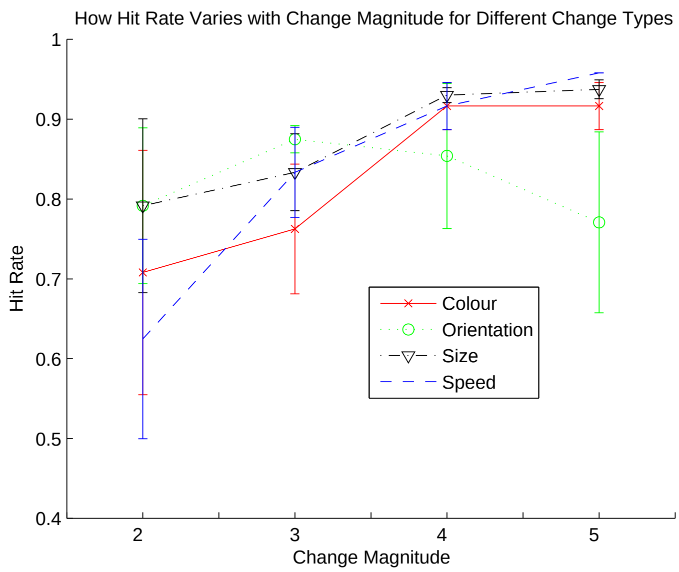


Figure 8.4: How hit rate varies with change magnitude for each of the four change types, averaged across participants.

8.1.3.3 Model Predictions

d' values were calculated for each of the average hit rates and false alarm rates used previously. These were then compared with d' values predicted by both a sample size model and a high-threshold model. Table 8.2 shows that regression coefficients were higher for the sample size model than the high-threshold model for all change types. However, every r^2 value was significant to $p < 0.05$. Figure 8.5 shows an example of the data for colour changes - d' plotted against $\frac{1}{\sqrt{N}}$ and $\frac{1}{N}$, and the regression fits to these data.

Table 8.2: Experiment 1: Regression Coefficients for obtained d' values vs. d' values predicted by two models of search behaviour

Feature Dimension	$r^2 for \frac{1}{\sqrt{N}}$	$r^2 for \frac{1}{N}$
Colour	0.95	0.73
Orientation	0.74	0.54
Size	0.90	0.65
Speed	0.77	0.49

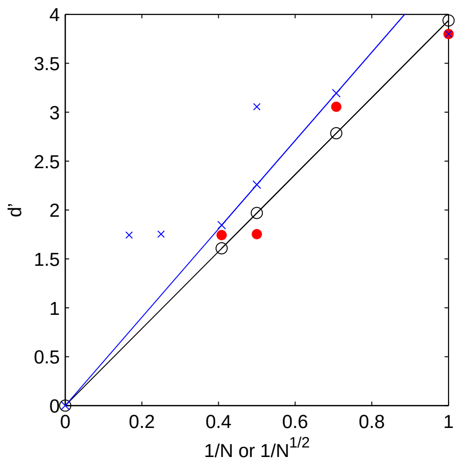


Figure 8.5: How well the sample size model and the high threshold model predict the data from Experiment 1. Red dots are $d' vs. \frac{1}{\sqrt{N}}$ and the corresponding regression fit is the black line. Blue crosses are $d' vs. \frac{1}{N}$ and the corresponding regression fit is the blue line.

8.1.3.4 Orientation Changes

Given that orientation changes exhibited normal set size effects, but an abnormal change magnitude effect, the effect of orientation changes on hit rate was examined for different magnitudes for each of the four different set sizes. This was done so that it could be determined whether the change magnitude effect shown in Figure 8.4 (i.e., a decline from magnitudes of 3 to 4 after an increase from 1 to 2) occurred across a limited range of set sizes, or all of them. The results of this analysis are shown in Figure 8.6, and show that this effect (an increase followed by a decrease) occurs only for a set size of 6. However, the other set sizes show no effect of change magnitude on hit rate, and this is why there is no overall effect of change magnitude for orientation.

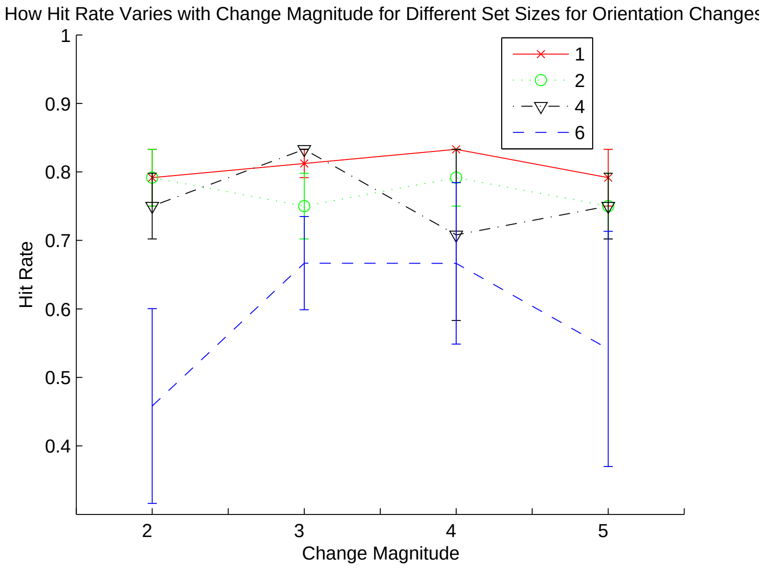


Figure 8.6: How hit rate varies with change magnitude for each of the four set sizes, for orientation changes. Data are averaged across participants.

8.1.4 Discussion

The prediction that changes to colour, orientation, size and speed would produce similar set size effects was supported (see Table 8.1). It was also predicted that data would conform better to a sample size model than a high-threshold model, and this too was supported. It was further predicted that changes in all stimulus variables would produce corresponding changes in performance. This prediction was supported for colour, size and speed but not for orientation. The lack of a consistent increase in performance with change magnitude for orientation changes suggests that increments in orientation do not produce corresponding increases in performance, in the same way that changes in the other stimulus dimensions do. It has been shown that orientation differences amongst neighbouring stimuli in a static search paradigm produce a reliable increase in performance with increases in the magnitude of difference (Bergen & Julesz, 1983). However, in a search-for-change paradigm, especially one in which multiple change types are used, there will be added complications. This paradigm may make differences between orientation and other stimulus variables more salient, and so make the detection of orientation changes much more difficult in comparison to detection of the other changes. The question is then, what differences are there?

Although the mathematical representation of an orientation change can be reduced to a scalar quantity (i.e., an increase in a certain number of degrees), orientation changes are not scalar in a physical sense - they involve rotation of an existing form. The changes to colour, size and speed are all physically scalar changes. Although speed is a vector quantity, the changes in speed in this experiment are never a change in direction, they are only increases in the drift rate of the gabor in the existing direction of motion. The discrimination of orientation changes could be further complicated by an automatic grouping when orientations 'line-up' or co-incide with one another (see Beck, 1966 for examples of this grouping effect). Another possibility is that detection of changes to the orientation of elements relied on processes at a higher level than detection of changes to other features. For instance, the envelope size of the gabors was relatively constant at different orientations and so the difference in orientation may only be apparent at the higher level of texture analysis (i.e., the elements appeared as oriented textures). Regardless of the reasons, it is clear that detecting changes in orientation is not as straight forward as detecting changes in the other stimulus dimensions. Therefore, orientation changes are not dealt with in most of the remaining experiments of this thesis.

8.2 Experiment 2 - Change to Stationary Gabor Stimuli Using Cues to Manipulate Set-Size

8.2.1 Introduction

In the last experiment, it was found that performance was not affected by the magnitude of orientation changes. It was suggested that this could be because orientation changes involve a more complex or 'higher-order' discrimination than do changes in the other stimulus dimensions or that similar orientations are grouped together. Orientation changes were excluded from this experiment so that only changes which produced a definite, positive relationship between performance and change magnitude were used. This relationship was required because this experiment was aiming to determine levels of change magnitude that were equivalent in terms of their detectability in different stimulus dimensions.

Another difference between this experiment and the last one is that instead of set size being manipulated by having a different number of elements present, it is manipulated by having the same number present on each trial, but cuing a different subset of those elements at the beginning of each trial. This is a *cued set size* or *relevant set size* (see Palmer et al., 1993) manipulation. It keeps the amount of sensory information relatively constant on each trial, and means the set size manipulation can be considered more as an attentional manipulation (i.e., manipulating attentional load), rather than a combined attentional and sensory manipulation. Furthermore, manipulating set size independently of the number of elements displayed keeps statistical decision noise constant (see Section 3.3.5.2) and so means results will not be contaminated by this effect. By eliminating another source of variability in the procedure, the experimental design adheres more closely to the psychophysical principles outlined in Chapter 3, Section 3.4.

8.2.2 Aims

- To determine the effect of changes in set size on the detection of a single element changing on a single dimension
- To determine a single set of conditions suitable for comparing single feature changes with multiple feature changes occurring simultaneously, across several objects and within the same object

8.2.3 Predictions

1. As for Experiment 1, it was predicted that set-size effects would be similar for colour, speed and size, in agreement with Palmer (1994) and Wilken and Ma (2004), and Experiment 1.

2. As for Experiment 1, it was predicted that increases in change magnitude would produce corresponding increases in performance (following Brown & Orbach, 1998) and, further, that the slope of the change mangitude-performance relationship would be similar for the different change types, due to the magnitudes of change being roughly equated in terms of detectability through pilot testing.

8.2.4 Method

8.2.4.1 Participants

Nine participants took part in this experiment, seven males and two females. All were volunteers recruited from the department of Human Movement Studies at the University of Queensland. Their mean age was 29.44 (range 21-42).

8.2.4.2 Procedure and Design

Trials in this experiment followed the same presentation paradigm for Experiment 1. An example trial is shown in Figure 8.7.

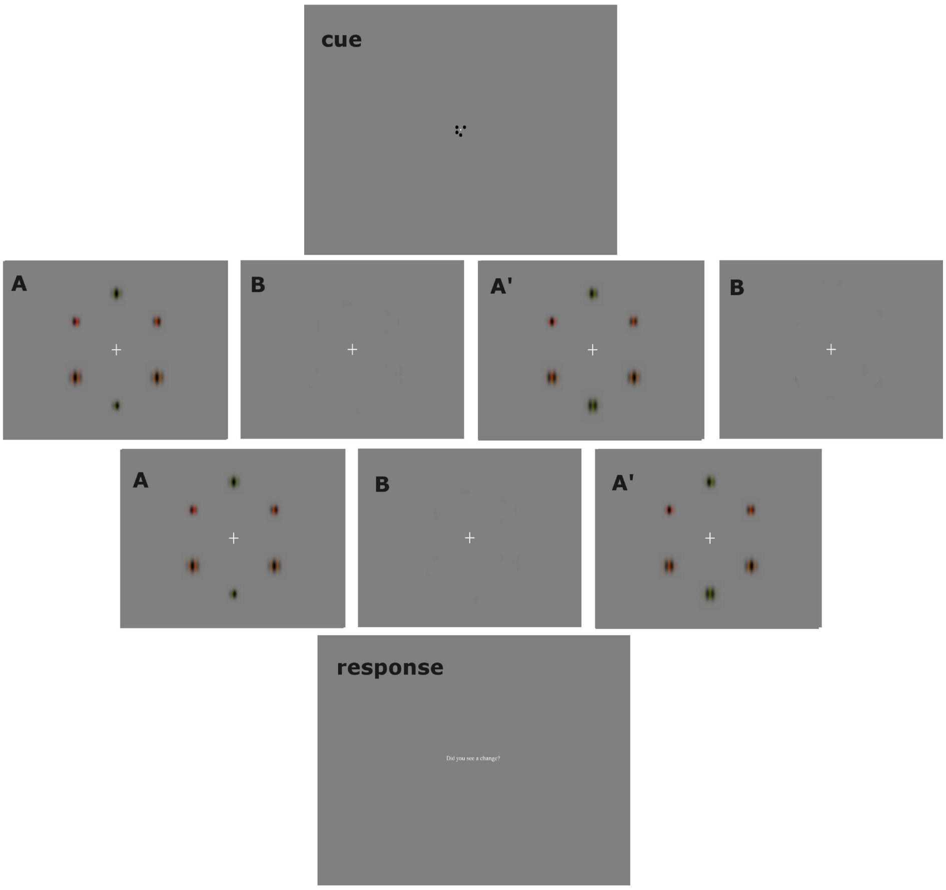


Figure 8.7: Example of a size change trial. The changing element is the one directly below fixation point (size change).

Independent variables in this experiment were the presence of change (present or absent), the type of change (colour, speed, size), the cued set-size (1, 2, 4 or 6), and the magnitude of change (1-4 steps). Therefore, there were 48 unique trials in which a change occurred and 48 in which no change occurred. These 96 trials were repeated 4 times in 4 separate blocks and the trials were run in a random order in each block.

8.2.4.3 Stimulus Configuration

Initial values of all elements and post-change values of changing elements were allocated in the same manner as for Experiment 1. The stimulus points are given below.

- Color (hue): 0.98, 1.01, 1.04, 1.08, 1.11, 1.21, 1.24, 1.27, 1.30, 1.32

- Size (degrees): 1.80, 1.98, 2.18, 2.40, 2.64
- Speed (degrees per second): 0.30, 0.42, 0.59, 0.82, 1.52, 1.61

The stimulus dimensions that did not undergo change - orientation, spatial frequency and luminance - were kept constant throughout (orientation: vertical, spatial frequency: 1.5 cycles per degree). Like Experiment 1, the values for size and speed lie on a logarithmic scale and follow Weber's law (see Equations 8.3, 8.4 and 8.5).

8.2.5 Results

Figures 8.8, 8.9 and 8.10 plot how performance (hit rate, see Section 8.1.3.1) varies with change magnitude and set size for colour, size and speed, respectively. These graphs show effects of set size and change magnitude for each change type.

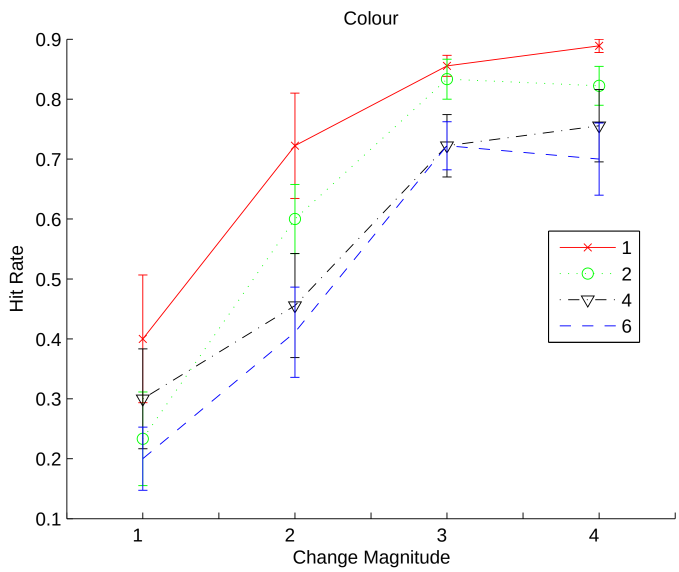


Figure 8.8: How performance changes with change magnitude and set size for colour changes.

Figure 8.11 shows how hit rate changes according to set size for each change type. The lines of best fit (from robust regression) and their standard errors are given in Table 8.3. A set size effect is present for all feature dimensions. The size of the effect is similar for size and speed and is smaller for colour.

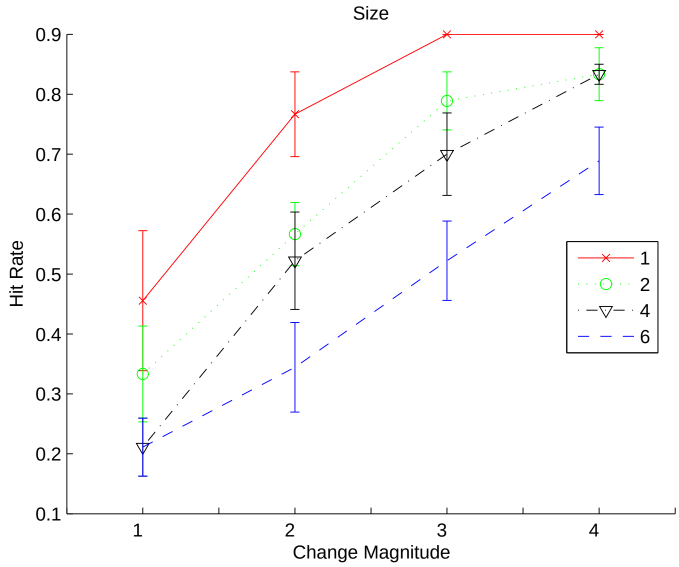


Figure 8.9: How performance changes with change magnitude and set size for size changes.

It is apparent from the data that performance was substantially below ceiling for set sizes of 1. This perhaps indicates that the cuing was not totally effective and/or the constant presence of other elements in addition to those being attended to created noise in the data (there were always six elements present).

Table 8.3: Experiment 2: Set-Size Effect Slopes and Error

Feature Dimension	Slope	Error
Colour	-0.0443	0.0103
Size	-0.0713	0.0113
Speed	-0.0621	0.0040

From Figure 8.11, it is clear that performance is most similar across change type for a set size of 4. Figure 8.12 includes data only from trials with a set size of 4 and shows how performance changes with change magnitude for the three different change types. The variability across different change types is similar for magnitudes of 2, 3 and 4.

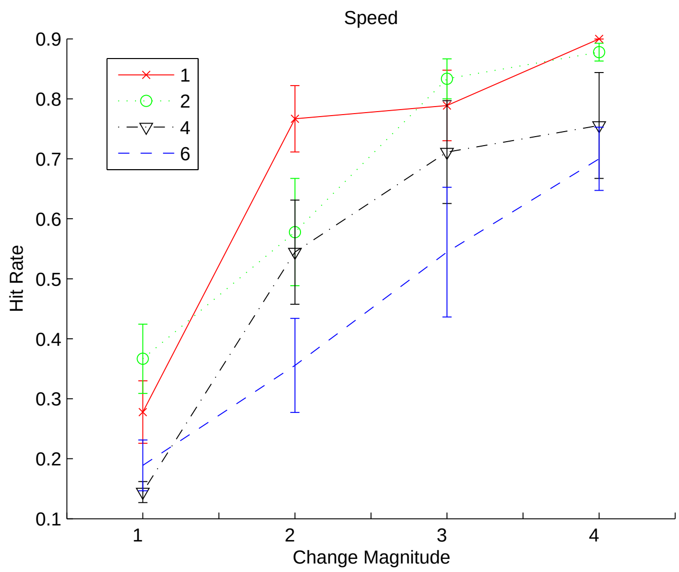


Figure 8.10: How performance changes with change magnitude and set size for speed changes.

8.2.6 Discussion

The predictions of similar set-size effects across change type was supported for size and speed, but not for colour (which yielded a smaller effect). Effects of change magnitude, however, were similar for colour, size and speed, in support of the results of Experiment 1 - this, however, was not quantified by slope calculation in the current experiment. The different set-size effects is something that requires further investigation - given that this was a more controlled paradigm than Experiment 1, these results may be more reliable. However, the results for set sizes of 1 were substantially below ceiling, whereas in Experiment 1, they were at ceiling. This indicates that the constant presence of distractors created noise, be it sensory or decision noise, that reduced performance for what should be an easily detectable change. In Palmer's (1998) experiments, it was found that both the display set size (i.e., no cue - the number of elements present is the same as the set size) and relevant set size manipulations produced ceiling effects at a set size of 1 and, furthermore, yielded very similar set size effects. The cue used in Palmer's studies however, was different to that used in the current experiment.

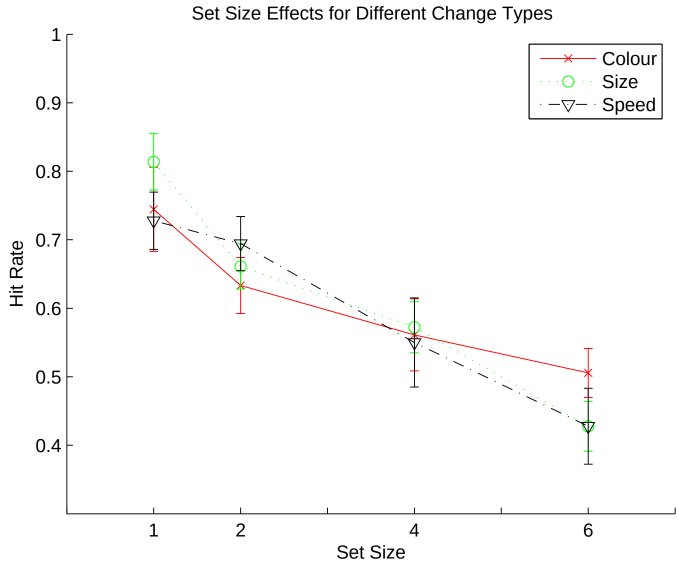


Figure 8.11: Set size effects for the different change types.

Palmer presented cues in the position of the elements, rather than at fixation (as in the current experiment). It is possible, therefore, that the cue presented at fixation is not as efficient as cues presented in element positions in terms of directing attention to those positions. This could create extra noise at both a sensory and a decision-based level in the visual system, as more noisy inputs are required to be integrated because they are not filtered out at a low enough level. Furthermore, Palmer measured set-size effects for individual observers and then averaged these, whereas in the current experiment, set size effects were generated from the average performance of all observers.

Given that the prediction of similar set-size effects across change type was not supported, it is possible that the similar set size effects obtained for simple features in standard visual search (see Palmer et al., 2000) does not hold for change detection experiments, because these experiments necessarily involve a transfer of stimulus information into memory, and processes such as comparison between stimulus elements may therefore occur at a higher level of the visual system. To further investigate this possibility, set size effects are again looked at in Experiment 4.

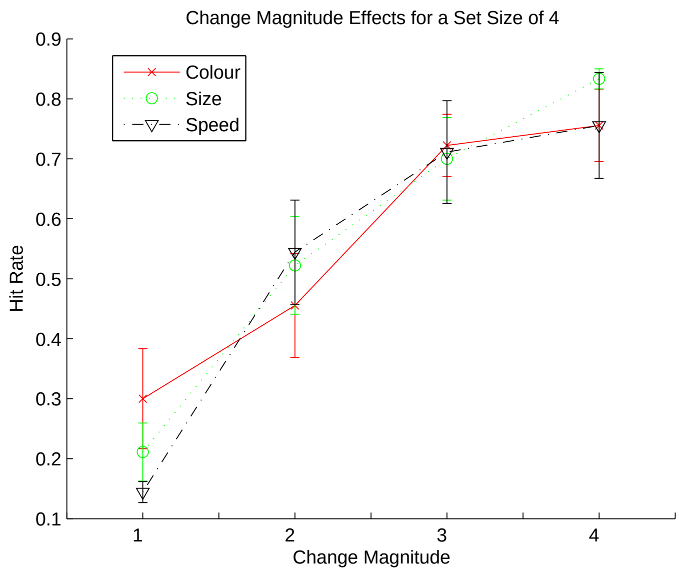


Figure 8.12: How hit rate varies with change magnitude for different change types when the set size is 4.

One of the aims of this experiment was to determine a set of conditions that produced equivalent levels of detectability across the three change types. The set size for which detectability across change type was most similar was 4. The equivalent levels of detectability were established so that a comparison can be made between single and multiple changes in the next experiment, where each stimulus increment of change is equivalent.

When multiple changes are made, it is possible that the overall detectability of this composite change will be additive. Therefore, it is important to choose a level of change magnitude that, as well as producing similar detectability across change type, produces a level of performance that is not so high such that it could produce ceiling effects in the multiple change situation (if the effects are indeed additive). Also, the ‘equivalence’ level of detectability should be substantially above chance (50%) so that a reliable baseline of detectability can be established in the single change case.

The combination of equivalence and being substantially below ceiling and above floor levels is fulfilled best by changes of magnitude 3 with a set size of 4 (see Figure 8.12). Variability of hit rate for these changes is very small and the mean hit rate is around .75, which is ideal as it is halfway between floor (chance) and ceiling.

Chapter 9

Multiple Changes

9.1 Experiment 3 - Multiple Changes Within and Between Objects

9.1.1 Aim

The aim of Experiment 3 was to look at interactions between different visual features (colour, size and speed), where the detectability of each individual change was relatively similar. In order to do this, a restricted set of conditions from Experiment 2 was employed (a constant set size and a constant magnitude of change on each dimension).

9.1.2 Predictions

It was predicted that:

1. Following Wilken (2001), performance would be determined by the number of features changing, rather than the number of objects changing (as in Luck & Vogel, 1997). Accordingly, it was expected that two features changing within an object would produce the same level of performance as two features changing across two objects.

2. Following Luck and Vogel (1997), it was expected that different features changing across two objects would produce the same level of performance as the same feature changing across two objects.

9.1.3 Participants

Ten participants took part in this experiment, four males and six females. All were undergraduate students at the University of Queensland and were paid $10 for their participation. Their mean age was 20.9 (range 18-24).

9.1.4 Procedure and Design

Each trial followed the presentation pattern used in Experiment 2. At the beginning of each trial, stimulus variables were configured in the same way as for Experiment 2. However, in this Experiment, more than one element and feature could change, but the magnitude of that change was fixed. The independent variables for this experiment were: number of objects changed (1-2), number of features changed (1-2) and feature type(s) changed (size, colour, speed). The experiment therefore followed a 2 (absence/presence of change) x 2 (number of objects changed) x 2 (number of features changed) x 3 (feature change type) design. When two objects changed, only one feature within each object would change. Therefore, when two objects changed, they could change on either the same feature (single feature change) or each change on a different feature (double feature change). The possible double feature change combinations were: colour & speed, speed & size and colour & size. Figure 9.1 shows an example of a trial in which one element changes size and one changes colour.

9.1.5 Results

Figure 9.2 shows the proportion of hits for the 1 and 2 object conditions where 1 and 2 features are involved. This data also shows a level of detectability that would be expected from probability summation alone - detectability will increase with two targets present simply because more suprathreshold targets are present. In other words, observers are more likely to spot a target (i.e., produce a ‘yes’ response) when more targets are present in the search field. The calculation of

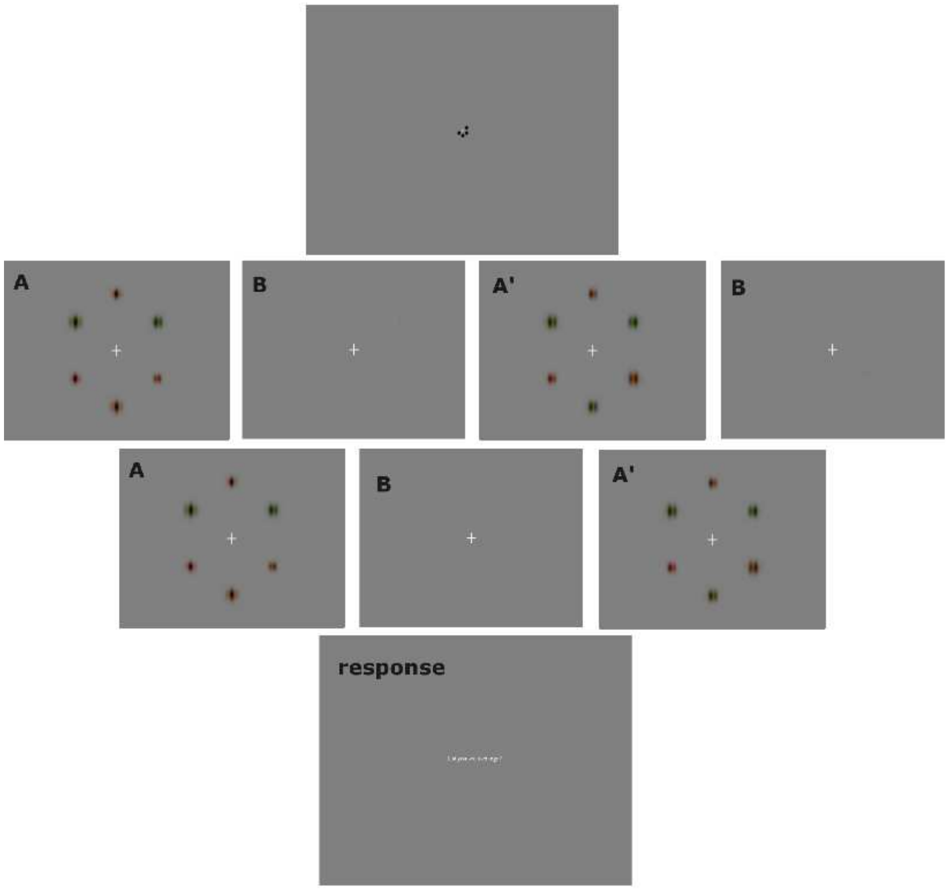


Figure 9.1: Example of a trial where one element changes size and one changes colour. One changing element is directly below fixation (colour) and one is diagonally down and to the right (size).

probability summation is given in Equation 9.1. This formula has been used in psychophysics detection studies to calculate the expected increase in detectability of a stimulus increment when two stimulus variables (e.g., contrast and orientation) are changing simultaneously (e.g., see Reisbeck & Gegenfurtner, 1998). It is also the formula used to calculate statistical decision noise from added distractors in visual search. The formula effectively calculates the probability of two independent channels both not detecting a signal, where the probability of one detecting a signal is p, the probability of it not detecting that signal is $1 - p$, the probability of both not detecting a signal is $(1 - p)^2$. Therefore, the probability of either of them detecting a signal is $1 - (1 - p)^2$.

$$p_{summ} = 1 - (1 - p)^n \tag{9.1}$$

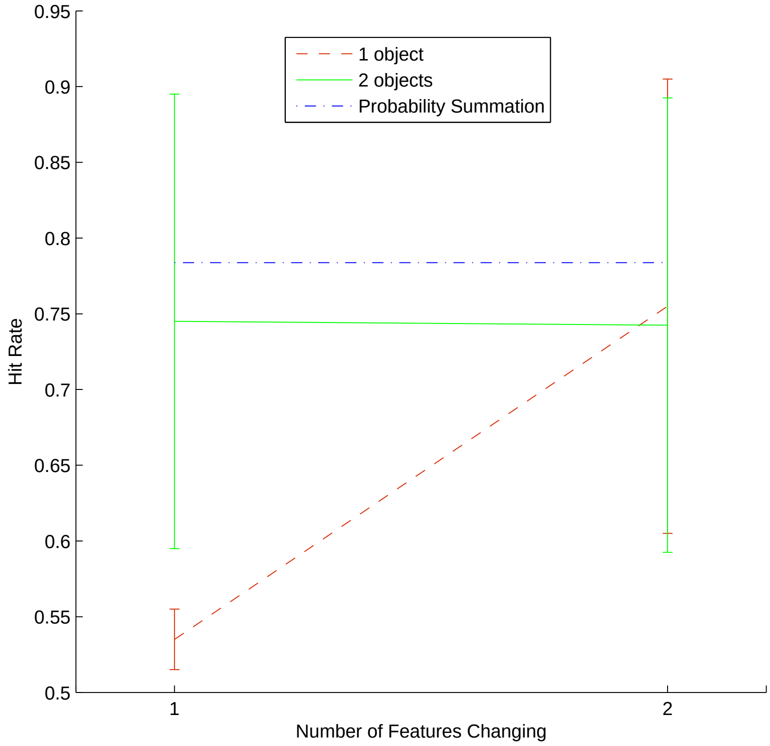


Figure 9.2: How the hit rate varies with the number of featural changes and number of objects changing. The level of performance expected from probability summation is shown.

Two important results are shown in Figure 9.2. Firstly, there is no significant difference between the two double feature conditions (i.e., when two features change within an object as compared to when they change across two different objects). Secondly, there is no significant difference between the probability summation level and all of the double object/feature conditions.

9.1.6 Discussion

The results indicate that multiple changes do not summate beyond the level expected from probability summation. This indicates that the detection of multiple changes can be accounted for by a simple model of parallel feature detection.

Furthermore, the lack of any difference between the two double feature conditions (two features changing within an object or across two objects) supports the prediction that performance would be determined by the number of features changing rather than the number of objects changing.

This result goes against the assertion of Luck and Vogel (1997) that performance in vSTM tasks is based on the number of objects changing, rather than the number of features. If this assertion were correct, we would expect a difference between conditions where changes occurred across objects and conditions where they occurred within a single object (an object-based benefit). However, the criticism of Luck and Vogel (1997)'s interpretation by Wheeler and Treisman (2002) and their results in favour of the feature-based interpretation of vSTM are more in agreement with the results of this experiment as is the results of Wilken (2001) Experiments 1-4 (see Section 4.4.5) - it is the number of features changing, rather than the number of objects, that dictates performance. It is possible, however, that the difference between two features changing across two objects and two features changing within a single object is small enough that it needs a more sensitive measure than the one provided by this experiment. Another possibility is that, for the current task, observers employed a serial-like search strategy in which only one object is searched at a time. If this were the case, multiple changes across objects would create no facilitation of detection performance, because only one of the changing objects was detected. Similarly, when a single object changes (regardless of whether it was on one or two features), it will be detected with the same level of performance as in the two-object change condition. However, it is important to note that multiple changes to a single object did improve performance substantially in comparison to single changes to a single object. This means that these changes somehow add to increase the salience of the single-object target. It is possible then, that observers are employing an object based search strategy, where multiple changes make a single object easier to detect, perhaps by improving the quality of its representation in vSTM.

For these reasons, multiple feature changes are looked at again in Experiment 12 after experiments 5-11 concentrate on developing more sensitive measures for change detection in a search paradigm.

Chapter 10

Set Size Effects Using a Left/Right Response Paradigm

10.1 Experiment 4 - Set-Size Effects Again

10.1.1 Introduction

This experiment involved a re-examination of the set-size effects for different change types. Set size effects were looked at in Experiments 1 and 2. These experiments developed a well-controlled paradigm using stationary gabor elements with moving bars, so that participants did not need to track moving elements. Also, in Experiment 2, the number of gabors was fixed at 6, and set size was manipulated using a cue, in a similar method to that used by Palmer (1994). This *cued set size* method meant that the amount of low-level sensory information present on any given trial was relatively similar, and that only the *attentional* information (given by the cue) was varied.

Although participants were instructed to fixate on the centre in Experiments 1 and 2, and there was no need for them to track elements, it is still possible their eyes moved away from fixation, to check changing elements, for example. Controls against eye movements are important as they can contaminate set size effects (see Huang & Pashler, 2005). Therefore, the current experiment employed a series of controls against eye movements. Participants were instructed to fixate throughout the experiment and were informed (correctly) that eye-movements would make the task more difficult and frustrating, even though they are intuitively appealing in search tasks. This experiment also employed a one-shot paradigm, in which each change was presented only once (rather than twice as in Experiments 1-2). This was done to further constrain the presentation and timing, to make eye-movement based search behaviour less likely. Another measure taken to reduce eye movements was the shortening of the presentation time of each individual display of the stimuli to 500 msec, which is just above the latency of a single eye movement. Also, a greater number of participants were used in this experiment, compared to Experiments 1-2. Finally, this experiment utilised a different response paradigm to Experiments 1-2. A change occurred on every trial and participants were required to identify whether it occurred on the right or left side of the screen. Therefore, responses in this experiment unambiguously consisted of hits and misses, without the presence of correct rejections and false positives. These modifications were introduced with the aim of reducing noise in the response data, not only due to eye movements, but also eccentricity effects and response bias and make any underlying effects more salient (i.e., increase statistical power).

10.1.2 Predictions

1. Set size effects were expected to produce log-log slopes consistent with the use of a limited capacity process in the detection of change, similar to the findings of Brown and Orbach (1998) and Palmer (1990)

10.1.3 Method

10.1.3.1 Participants

Twenty participants took part in the experiment. All were students at the University of Queensland, recruited through an internet advertisement and each was paid $10 for their participation. Nine were female and eleven were male. Their ages ranged from 18 to 43 and the mean age was 23.75.

10.1.3.2 Procedure and Design

Trials in this experiment followed a one-shot paradigm:

$$cue \rightarrow A \rightarrow B \rightarrow A' \rightarrow response,$$

where the cue was presented for 1000 msec, A and A′ for 500 msec each and B for 120 msec. The response screen was presented until the participant responded.

This experiment followed a 4 (cued set size) x 5 (change type) x 5 (change magnitude) x 2 (side of target) design. The 250 trials created were randomly arranged in each of 5 blocks. For the stimulus dimension on which the change occurred, half of the elements were given the lowest point on that dimension and the other half were given the point to which the target would change (following the method of Rensink, 2000d). For each of the other dimensions, half of the elements were allocated the lowest point on that dimension and the other half were allocated a point randomly selected from the pre-set points (see Table 10.1). This experiment included orientation despite the finding from Experiment 1 that there was no effect of change magnitude for orientation. This was because the greater control employed in this experiment could potentially reveal an effect of orientation change where the less controlled Experiment 1 didn't. For instance, this experiment used a shorter presentation time (500 msec) than Experiment 1 and also used a one-shot rather than a two-shot paradigm, and it is possible that whatever was producing the lack of change magnitude effect for orientation (e.g., grouping) relies on the greater time of stimulus presentation or some other aspect that was less controlled in Experiment 1.

10.1.4 Results

Figures 10.1, 10.2, 10.3, 10.4 and 10.5 show how performance varies with change magnitude for different set sizes for colour, orientation, size, spatial frequency

Table 10.1: Experiment 4: Stimulus Parameters.

Feature Dimension	min	max	step size
Orientation (°)	0	84	6
Size (°)	1.80	3.2	0.1
Spatial Frequency (cpd)	1.5	4.3	0.2
Speed (° s^{-1})	0.30	1.14	0.07

and speed changes respectively. Table 10.2 lists the slope and error values for the best fitting lines (from robust regression) for the magnitude vs. hit rate data collapsed over set size.

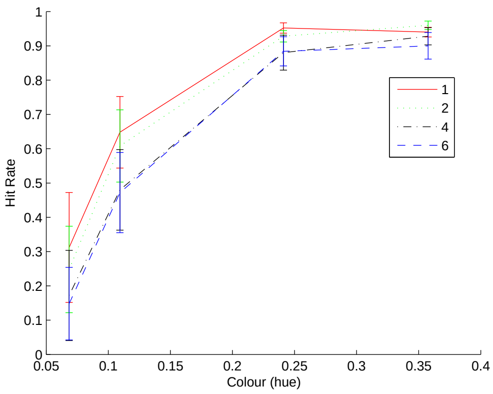


Figure 10.1: How performance for detecting colour changes varies with change magnitude for different set sizes.

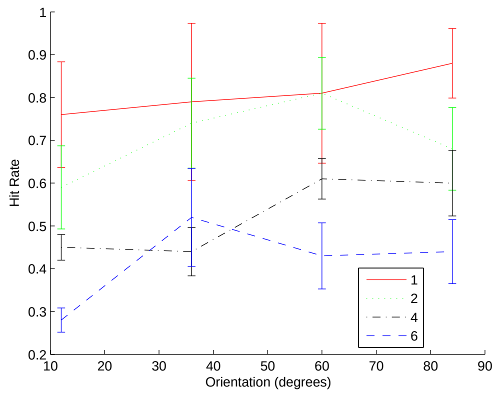


Figure 10.2: How performance for detecting orientation changes varies with change magnitude for different set sizes.

Table 10.2: Experiment 4: Magnitude Effect Slopes and Error

Feature Dimension	Slope	Error
Colour	2.3205	0.8531
Orientation	0.0018	0.0008
Size	0.334	0.0782
Speed	0.1584	0.0387
Spatial Frequency	0.133	0.048

The slopes in Table 10.2 are not comparable with one another, as they depend on the scale of each stimulus dimension. However, they can indicate trends within each stimulus dimension. For instance, it appears that hit rate does not increase

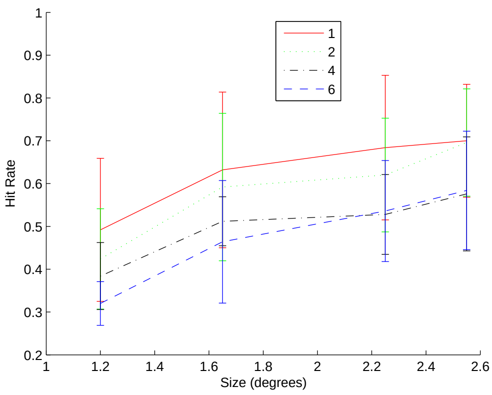


Figure 10.3: How performance for detecting size changes varies with change magnitude for different set sizes.

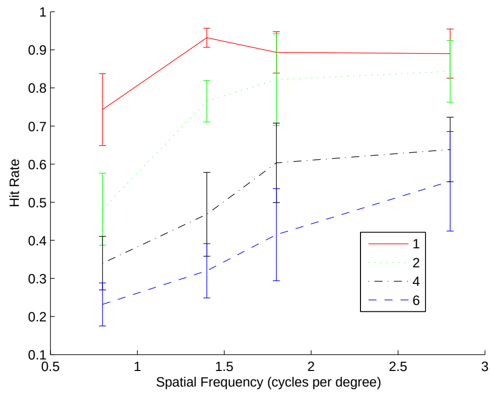


Figure 10.4: How performance for detecting spatial frequency changes varies with change magnitude for different set sizes.

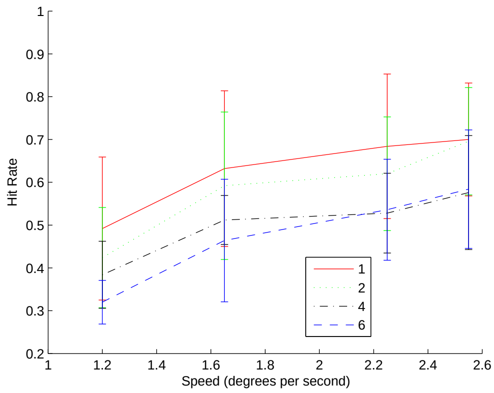


Figure 10.5: How performance for detecting speed changes varies with change magnitude for different set sizes.

with increases in the magnitude of orientation changes. This was also found in Experiment 1 and shows this effect is present even in more controlled conditions used in this experiment (i.e., a one shot paradigm, same change magnitudes having the same physical size, shorter presentation times).

Figures 10.6, 10.7, 10.8, 10.9 and 10.10 show how performance varies for different set sizes for colour, orientation, size, spatial frequency and speed changes respectively. Dotted curves are for individual observers while the bold curve is the average of these. Table 10.3 gives the slopes and errors for the best fitting lines (yielded by robust regression) to the set size data. These data show, like Experiment 2, that performance was below ceiling for a set size of 1, in each case except perhaps for spatial frequency.

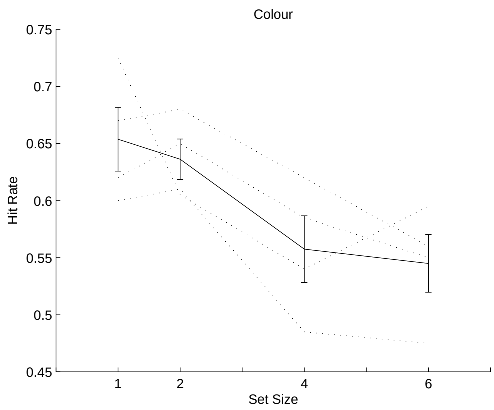


Figure 10.6: How performance for detecting colour changes varies with change magnitude for different set sizes.

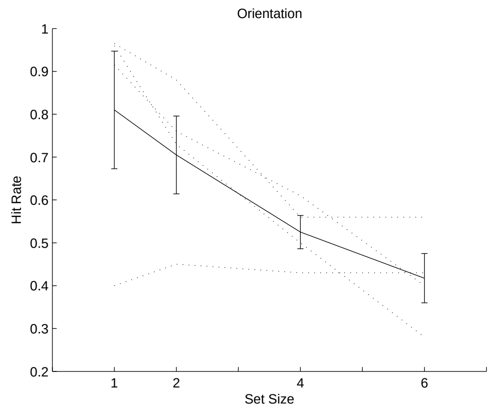


Figure 10.7: How performance for detecting orientation changes varies with change magnitude for different set sizes.

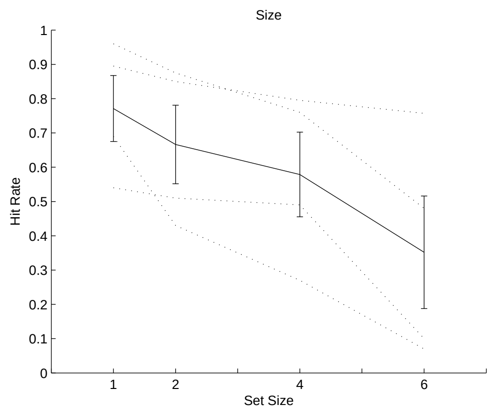


Figure 10.8: How performance for detecting size changes varies with change magnitude for different set sizes.

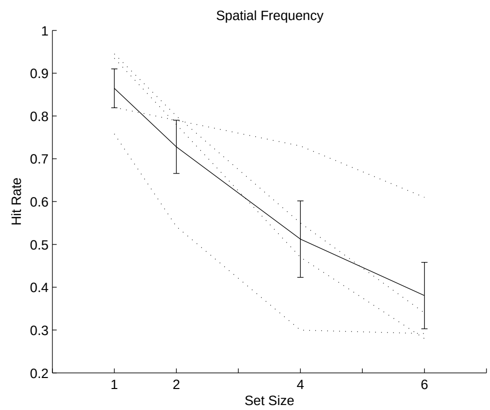


Figure 10.9: How performance for detecting spatial frequency changes varies with change magnitude for different set sizes.

Using the method of Palmer et al. (2000), data were transformed to log-log coordinates of threshold (taken as 0.5) vs. set size for each stimulus dimension. The slopes obtained from these plots are given in Table 10.4. Slopes are comparable for colour, size and spatial frequency, but the slope for speed is much higher. The slopes for colour, size and spatial frequency would tend to indicate unlimited capacity processing, while the slope for speed would indicate some form of limited capacity (see Section 3.4.3.2).

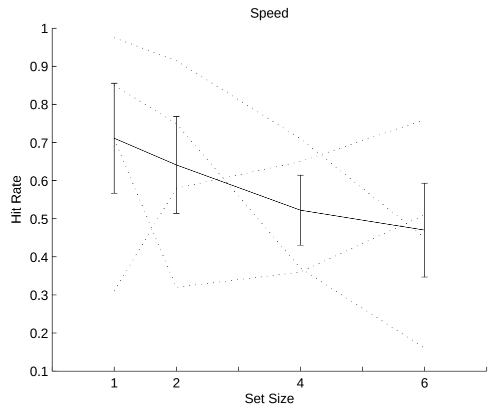


Figure 10.10: How performance for detecting speed changes varies with change magnitude for different set sizes.

Table 10.3: Experiment 4: Set-Size Effect Slopes and Error

Feature Dimension	Slope	Error
Colour	-0.0237	0.0054
Orientation	-0.0789	0.0082
Size	-0.0791	0.0116
Speed	-0.0487	0.0072
Spatial Frequency	-0.0967	0.0107

10.1.5 Discussion

The fact that no change magnitude effect was produced for orientation in this experiment shows that the effect demonstrated in Experiment 1 is robust to the modifications made to the procedure to make it more controlled (i.e., using a constant set size, using cues, using shorter presentation times and using a one-shot paradigm). In this experiment, the 'grouping effect' discussed previously should be equal in each trial (due to there being an equal number of elements present), unless attention filters out the gabors flanking the targets at a low enough level. The presence of a set size effect for orientation suggests that the distracting gabors are successfully filtered out.

This gives greater credence to the points raised in the Discussion of Experiment 1 relating to the differences between orientation and other stimulus variables. However, it does not validate them completely, as there could be other

Table 10.4: Experiment 4: Log-Log Set Size Effect Slopes and r^2 Values

Feature Dimension	Slope	r^2
Colour	0.2293	0.93
Size	0.3046	0.96
Speed	0.9655	0.97
Spatial Frequency	0.2895	0.96

reasons why orientation changes produce different response behaviour compared with changes in size, speed and colour. The next experiment investigates the different stimulus variables in a conventional search paradigm, where the target element does not undergo a change, but is different from the distractors on some stimulus dimension in the single presentation that occurs in each trial. Given that visual search studies have found that orientation targets are discriminable from distractors, and, furthermore, that increasing the difference in orientation between the target and distractors reliably increases performance, using this conventional search paradigm with the current stimuli and the current conditions should give us greater insight into the differences between orientation and the other stimulus variables.

Another interesting finding of the current experiment was that of similar set-size vs. threshold log-log slopes for colour, size and spatial frequency, with a much higher slope for speed changes. This goes against the prediction of similar log-log slopes for all change types which reflected limited capacity processing. The slope values for colour, size and spatial frequency would tend to suggest that detection of these changes, at least in the current task, is subject to an unlimited capacity process (see Palmer et al., 2000). The higher value for speed changes, however, would tend to suggest these changes are subject to a limited capacity process. The low slope values for colour, size and speed are perhaps surprising, given the finding of Brown and Orbach (1998) of a slope of around 0.75 for detection of contrast changes. Furthermore, given that Palmer (1990) found slopes around 0.5 as being consistent with a visual memory task (judging the line length of a test stimulus against a previously presented stimulus), it would perhaps be expected that the slope values in the current task be closer to 0.5 or 0.75. However, the slopes in the current experiment were taken from set size data that had been averaged across observers and not fitted to any function. A better method to use would be to fit each individual's set size data (stimulus magnitude vs. performance for each set size) to a psychometric function, take the change magnitude value at a fixed level of performance and so create a log-log slope for each individual observer and then average the slope values to get a more precise estimate of log-log set size vs. threshold slope (see Palmer, 1998). Furthermore, the task of Brown and Orbach (1998) involved identifying which of two different displays contained an element with a contrast increment, whereas this task involved an increment always being present on the second stimulus presentation (i.e., A′) within a trial. Perhaps then the current task was less difficult and subject to a different processing style. Another consideration relates to the effectiveness of the cues in the current experiment. Given that set size effects did not yield performance at ceiling for set sizes of 1, it is apparent that the constant presence of the 6 elements in the display is creating noise in the performance data even though attention *should* be directed to only a single element (assuming the cue is maximally effective). Unfortunately, the presence of this noise means that the relevant models may not be able to predict performance in these conditions. The presence of this noise would be more constant across set size than predicted by the models, which would yield shallower set-size vs. threshold slopes. Therefore, it may not be meaningful to compare data from the current experiment with the predictions made based by these models.

Chapter 11

Threshold Measurement using a Left/Right 2AFC Paradigm

11.1 Experiment 5 - Thresholds for Non-Changing Targets

11.1.1 Introduction

The aim of Experiment 5 was to establish baseline thresholds for several participants for the detection of targets defined by colour, size, orientation, spatial frequency and speed. This experiment did not involve a stimulus changing like previous experiments, but instead involved a single presentation in which the target was defined by a difference from all the other homogenous distractors. Establishing thresholds for each observer allows later experiments to be tailored to individuals, by having the change magnitudes scaled relative to individuals' thresholds. This should allow a more robust and accurate interpretation of the data. Given that the experimental procedure is well-controlled and adheres well to the principles of psychophysical measurement outlined in Section 3.4, it is ideal for measuring thresholds of individuals for each change type.

11.1.2 Method

11.1.2.1 Participants

Four participants took part in the experiment. All were males recruited from the Human Movements Laboratory and the Psychology Department at the University of Queensland, with ages ranging from 22-32 (mean age 26). Two participants were paid $10 for their participation.

11.1.2.2 Procedure

Figure 11.1 shows an example of the stimulus presentation. The fixation cross was displayed for 1000 msec before the stimulus presentation, which was displayed for 500 msec. The shortened presentation time was used to eliminate, or at least reduce, eye movements (see Experiment 5). The prompt was presented until the participant made a response. Because there was no interval over which the target changed, this was a standard visual search task, rather than a visual search-for-change task. On each trial, a target was presented. In half of the trials, the target was on the left and in the other half, it was on the right. Participants were required to indicate which side the target was on. As for Experiment 5, this left/right response paradigm was used to reduce ambiguity in the response data that could arise from yes/no response bias.

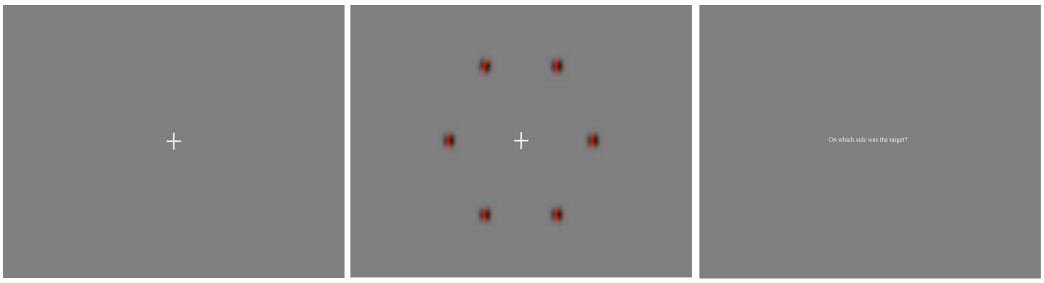


Figure 11.1: Example of a trial with an orientation target.

11.1.2.3 Conditions

This experiment followed a 2 (target-side) x 5 (target type) x 7 (target increment) design. Table 11.1 shows the minimum and maximum values for orientation, size, speed and spatial frequency as well as the step sizes which defined the steps in between. The colour hue values (h in L*c*h, CIE 1976) are listed below the table. Orientation is included again as a stimulus variable in this experiment, because the task is completely different to the change detection tasks in the previous experiments, and so there is no reason to assert that orientation targets will not produce a meaningful set of results in the current experiment.

Unlike Experiments 1-3, the step sizes are absolute - none of the feature dimensions are logarithmically scaled. This is so that the size of each step, and the size of the range between the lowest and highest steps are both reduced. This reduction is necessary because the detection of a target in this experiment is much easier than in the previous experiments. A single target was present on each trial and took a value of one of the steps on the feature dimension defining it (target type), but had minimum values on all other dimensions. All the distractor elements had minimum values (shown in Table 11.7) on all feature dimensions. The only difference amongst distractors was the direction of their motion (left or right, chosen at random). Therefore, the target was a feature singleton as defined by H. E. Pashler (1998).

11.1.2.4 Stimuli

Table 11.1: Experiment 5: Stimulus Parameters for Size and Speed.

Feature Dimension	min	max	step size
Orientation (°)	0	42	3
Size (°)	1.80	2.45	0.05
Spatial Frequency (cpd)	1.5	2.9	0.1
Speed ($^\circ$ s^{-1})	0.30	0.755	0.035

Colour (hue): 0.967, 0.980, 0.997, 1.007, 1.036, 1.044, 1.067, 1.077, 1.099, 1.109, 1.208, 1.242, 1.265, 1.301, 1.325.

11.1.3 Results

To determine if participants were responding preferentially to one side of the screen (left or right), response bias was estimated using the c values obtained from a signal detection analysis applied to all the data for each participant (see Table 11.8). c is measured in standard deviation units and so we can use 1 standard deviation as a criterion for excluding data. Given that all values were within one standard deviation of 0 (no response bias), all data were kept (see Table 11.2).

Table 11.2: Experiment 5: c Values for Participants from SDT Analysis

	AB	TW	SC	WM
c	0.01	-0.03	0.04	0.22

Psychometric function curves were plotted for each observer for each target type. These curves were constructed using the *psignifit* package, which is an addon to MATLAB 7.0. The curves constructed were of the default type, logistic, and can be described by the equation:

$$F(x) = \frac{1}{1 + e^{(a-x)/b}} \tag{11.1}$$

, where a and b are scaling parameters. Figure 11.2 shows an example curve for observer SC for speed targets and Figure 11.3 shows one for observer AB for colour targets. 75% correct thresholds were taken from each curve as was the slope at the 75% point (i.e., the derivative of the psychometric function where $F(x) = 0.75$). Threshold and slope values for the three included observers are listed in Tables 11.3 and 11.5, respectively. Included in Tables 11.4 and 11.6 are the values obtained by the BCa bootstrap method for one standard deviation below and above the mean. BCa was used as it has been found to be the most stable method of estimating confidence limits for thresholds and slopes (Wichmann & Hill, 2001b).

For a given change type, thresholds are relatively similar across participants. The same can be said of the slope values, but the large confidence intervals suggest this comparison should be made with caution. However, an obvious trend does emerge when comparing slope values between change types. Colour and speed changes yielded similarly high slope values, while size and spatial frequency produced similar, lower values and orientation targets gave a very shallow slope.

Table 11.3: Experiment 5: Thresholds

Feature Dimension	AB	TW	SC	WM
Colour (h)	1.02	1.03	1.01	1.01
Orientation (°)	6.66	7.01	8.83	5.62
Size (°)	1.89	1.90	1.98	1.93
Spatial Frequency (cpd)	1.74	1.78	1.94	1.88
Speed ($^\circ$ s^{-1})	0.34	0.34	0.42	0.38

Table 11.4: Experiment 5: Thresholds Confidence Intervals

Feature Dimension	AB	TW	SC	WM
Colour (h)	1.01, 1.03	1.02, 1.04	1.01, 1.02	1.01, 1.02
Orientation (°)	5.62, 7.50	5.83, 7.92	7.95, 9.23	4.36, 6.48
Size (°)	1.87, 1.90	1.88, 1.91	1.95, 1.99	1.92, 1.95
Spatial Frequency (cpd)	1.71, 1.77	1.74, 1.80	1.88, 1.97	1.77, 1.92
Speed ($^\circ$ s^{-1})	0.33, 0.35	0.33, 0.35	0.40, 0.44	0.37, 0.39

Table 11.5: Experiment 5: Slopes

Feature Dimension	AB	TW	SC	WM
Colour (h)	13.16	9.29	12.32	16.76
Orientation (°)	0.11	0.09	0.20	0.08
Size (°)	6.00	4.95	3.42	5.02
Spatial Frequency (cpd)	3.85	3.44	1.70	0.93
Speed ($^\circ$ s^{-1})	13.13	15.07	3.07	10.57

Table 11.6: Experiment 5: Slopes Confidence Intervals

Feature Dimension	AB	TW	SC	WM
Colour (h)	10.62, 18.97	6.74, 10.58	9.33, 14.36	12.69, 21.73
Orientation (°)	0.08, 0.18	0.06, 0.12	0.13, 0.45	0.06, 0.11
Size (°)	4.17, 9.24	3.51, 6.83	2.71, 4.61	3.76, 6.59
Spatial Frequency (cpd)	2.71, 5.84	2.43, 4.84	1.34, 2.32	0.74, 1.31
Speed ($^\circ$ s^{-1})	7.72, 19.27	9.36, 22.07	2.04, 3.53	7.32, 15.58

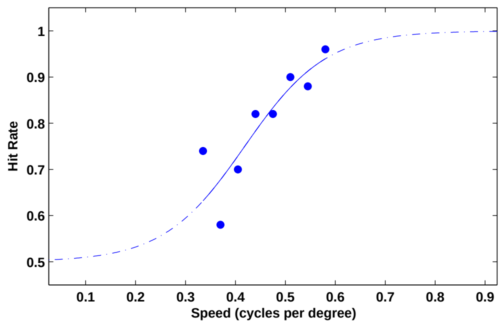


Figure 11.2: A psychometric curve plotted for speed targets for participant SC.

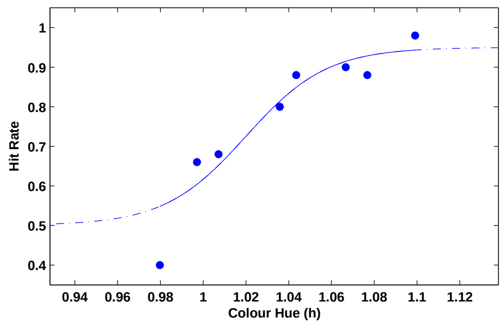


Figure 11.3: A psychometric curve plotted for orientation targets for participant AB.

11.1.4 Discussion

Given that slope values reflect the rate of change in performance with respect to the change in stimulus values, it is likely that the consistent differences in slope values between change types reflect, to some extent at least, intrinsic differences in the processing of different stimulus dimensions. However, these slope values are still tied to the scaling of the different dimensions.

The very shallow slope values for orientation reinforce the findings of Experiments 1 and 4 that orientation changes do not produce a reliable increase in performance with an increase in the magnitude of orientation change. In this experiment, the targets were defined by a difference across space and not a change over time, and so this indicates that orientation discrimination itself, and not just detection of orientation changes, is a more difficult task compared to discrimination of other stimulus variables (at least in the current stimulus conditions). Although other studies have found reliable effects of orientation target-distractor difference (see Bergen & Julesz, 1983), these studies did not involve discrimination of multiple stimulus variables within the same experiment, and so the participant would have been better attuned to pick up the differences in orientation. Furthermore, the Bergen and Julesz (1983) study used elements that were closely packed together and the difference between the target and homogenous distractors would therefore be more obvious (see Figure 11.4).

Figure 11.4: Example of a trial from the orientation discrimination/search experiment of Bergen and Julesz (1983).

11.2 Experiment 6 - Thresholds for Changing Targets

11.2.1 Introduction

This experiment used the same 2AFC left/right response paradigm as the last experiment, but in each trial, the target was defined by a change and a one-shot paradigm was used, as in Experiment 4, so that the change was presented once in each trial.

11.2.2 Method

11.2.2.1 Procedure

The presentation sequence and timing was the same as for Experiment 4:

$$cue \rightarrow A \rightarrow B \rightarrow A' \rightarrow response$$

The cue lasts 1000 msec, A and A' last 500 msec each and B lasts 120 msec, while the response screen is present until the observer responds. In each display of A, half (3) of the distractors had the pre-change value on the target stimulus dimension and half (3) had the post-change value. Therefore, in each A' display, half + 1 (4) had the post-change value and half - 1 (2) had the pre-change value. This was done so that the target could be discriminated from distractors on the basis of its change over time only - it could not be detected simply on the basis of its appearance on the post-change display.

11.2.2.2 Participants

Three participants took part in the experiment. All were males recruited from the Human Movements Laboratory and the Psychology Department at the University of Queensland, with ages ranging from 22-32, and an average age of 25. Two participants were paid $10 for their participation.

11.2.2.3 Conditions

This experiment followed a 2 (target side) x 5 (change type) x 7 (change magnitude) design. Changes only occurred from the minimum point on the featural dimension along which the change took place. Therefore, a change of a given magnitude always had the same physical magnitude for a particular featural dimension. Table 11.7 shows the minimum and maximum values for orientation, size, speed and spatial frequency as well as the step sizes which defined the steps in between. As for Experiments 4 and 5, no featural dimensions were logarithmically scaled. The colour hue values (h in L*c*h, CIE 1976) are listed below the table. For all dimensions, the step sizes and maximum values are larger than those used in Experiment 5 because the task is harder (change detection amongst heterogenous distractors as opposed to search for a feature singleton amongst homogenous distractors).

Colour Values (hue): 0.967, 0.997, 1.036, 1.067, 1.099, 1.208, 1.265, 1.325.

Table 11.7: Experiment 6: Stimulus Parameters for Size and Speed.

Feature Dimension	min	max	step size
Orientation ($^\circ$)	0	84	6
Size ($^\circ$)	1.80	2.6	0.1
Spatial Frequency (cpd)	1.5	3.1	0.2
Speed ($^\circ$ s^{-1})	0.30	0.86	0.07

11.2.3 Results

As for Experiment 5, c values were calculated to identify any response bias the participants had (see Table 11.8). These values are given below. No c values were above 1 or below -1 and so all participants' data were retained for further analysis.

Psychometric function curves were plotted for each observer for each target type, again using the *psignifit* package. Figure 11.5 shows an example curve

Table 11.8: Experiment 6: c Values for Participants from SDT Analysis

	AB	TW	WM
c	-0.10	-0.12	0.162

Table 11.9: Experiment 6: Thresholds

Feature Dimension	AB	TW	WM
Colour (h)	1.06	1.11	1.06
Orientation (°)	21.97	50.87	171.89
Size (°)	2.17	2.08	2.25
Spatial Frequency (cpd)	2.65	2.35	2.87
Speed (° s^{-1})	0.55	0.60	0.84

for observer AB for orientation targets and Figure 11.6 shows an example curve for observer TW for spatial frequency targets. 75% correct thresholds were taken from each curve as was the slope at the 75% point. Threshold and slope values, as well as 85% upper and lower confidence bounds, for the three included observers are listed in Tables 11.9, 11.10, 11.11, 11.12.

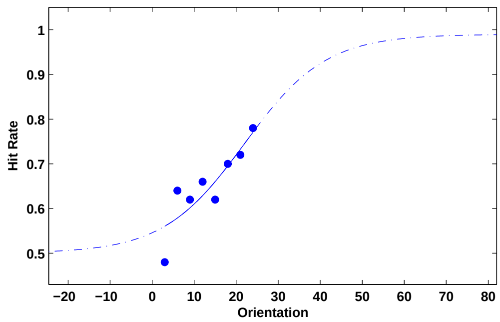


Figure 11.5: A psychometric curve plotted for orientation targets for participant AB.

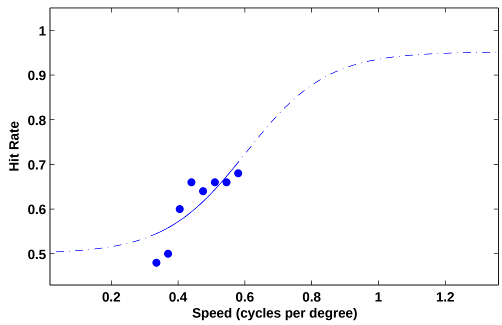


Figure 11.6: A psychometric curve plotted for spatial frequency targets for participant TW.

The threshold values are relatively similar across participants for all change types except orientation. However, it is clear that the orientation data for participant WM is invalid, as the threshold is very high (far higher than the maximum value tested) and the slope is very close to zero.

Table 11.10: Experiment 6: Threshold Confidence Intervals

Feature Dimension	AB	TW	WM
Colour (h)	1.05, 1.07	1.10, 1.16	1.05, 1.06
Orientation (°)	19.72, 27.13	29.65, 1141.9	38.37, 13496
Size (°)	2.14, 2.19	2.07, 2.13	2.19, 2.44
Spatial Frequency (cpd)	2.37, 4.19	2.31, 2.53	2.36, 3.83
Speed ($^\circ$ s^{-1})	0.53, 0.57	0.56, 0.75	0.63, 14.24

Table 11.11: Experiment 6: Slopes

Feature Dimension	AB	TW	WM
Colour (h)	6.68	11.97	11.81
Orientation (°)	0.03	0	0
Size (°)	3.66	1.97	1.55
Spatial Frequency (cpd)	0.60	4.48	0.92
Speed ($^\circ$ s^{-1})	4.47	2.06	1.62

Table 11.12: Experiment 6: Slope Confidence Intervals

Feature Dimension	AB	TW	WM
Colour (h)	4.52, 8.21	5.00, 22.61	9.20, 16.81
Orientation (°)	0.02, 0.04	0, 0.02	0, 0.01
Size (°)	2.38, 5.55	1.32, 2.46	0.75, 2.59
Spatial Frequency (cpd)	0.17, 1.14	1.24, 20.49	0, 1.83
Speed ($^\circ$ s^{-1})	2.89, 6.51	0.98, 3.14	0, 3.65

In comparison to Experiment 5, the thresholds in this experiment are higher overall while the slopes are lower. The higher thresholds would be expected given that the task is more difficult - there is more distracting information (heterogenous distractors) and participants have to detect a change over time rather than a salient target in a single presentation. The smaller slopes reflect the fact that more stimulus information is required for the same increment in performance - this would also result from increased task difficulty.

11.2.4 Discussion

Psychometric function slopes are a way of describing the relationship between detectability and stimulus intensity for any detection task. The slopes for colour and orientation in this experiment are similar to those in the last experiment (involving no change), while the size, spatial frequency and speed slopes are smaller in this one. However, in both experiments there is less consistency across participants in the size, spatial frequency and speed slopes. The extra variability in these dimensions makes differences between them less interpretable. To reduce the variability of response data within each stimulus dimension and so make the differences between change types more interpretable, it is necessary to use a more precise method of threshold establishment for changing targets. This is done in the next experiment.

11.3 Experiment 7 - Adaptive Thresholds for Changing Targets

11.3.1 Introduction

There was great variability in the threshold and slope measurements obtained in Experiment 6. The psychometric functions and their related statistics were calculated from a range of values that changed dramatically across participants. Therefore, to obtain more reliable and accurate psychometric functions and related statistics, this experiment tailors the range of tested stimulus values to the individual participant. This is done using a method known as *adaptive testing*, where the stimulus values presented to an observer are changed depending on their response history.

11.3.1.1 The PEST Algorithm

A popular and reliable method of obtaining thresholds adaptively is Parameter Estimation by Sequential Testing (PEST), introduced by Taylor and Creelman (1967) for the purposes of establishing auditory detection thresholds. The algorithm is governed by two basic rules which determine when to change the stimulus level, and, in the event the stimulus level is changed, what stimulus level to change to.

11.3.1.2 When to Change Levels in PEST

After each trial, permissible upper and lower levels are calculated for the number of correct trials, N(C), by first calculating the expected number correct, E[N(C)], using the level of performance corresponding to the threshold, P_t and the number of trials, T (see Equation 11.2). At the beginning of each trial, N(C) is compared to the upper and lower bounds for N(C), N_b(C), given by Equation 11.3.

$$E[N(C)] = P_t * T \tag{11.2}$$

$$N_b(C) = E[N(C)] \pm W \tag{11.3}$$

W is a constant called the *deviation limit*, which affects the speed with which the algorithm converges on a threshold stimulus value and the reliability of this value (faster convergence means less reliability and vice versa). Taylor and Creelman (1967) suggest a value of 1.5-2 for W in 2AFC experiments aiming for a 75% threshold.

If N(C) lies outside the bounds calculated by Equation 11.3, the stimulus level is changed, moving higher if N(C) is below the lower bound and moving lower if N(C) is above the upper bound.

11.3.1.3 What Stimulus Level to Try Next in PEST

There are no strict rules regarding choice of the first level - Taylor and Creelman (1967) suggest the choice can be casual, but should be a level that is easily detectable by the observer. For subsequent steps, the size is modified according to the following rules:

1. On every reversal of step direction, halve the step size
2. The second step in a given direction, if required, is the same size as the first
3. The fourth and subsequent steps in a given direction are each double their predecessor
4. For the third step in a given direction:
    - If the step immediately preceding the last reversal resulted from a doubling, the third step is not doubled.
    - If the step before the last reversal was not double its predecessor, the third step is double the second

11.3.1.4 Modifications to the PEST procedure

The current experiment used a modified version of the PEST procedure, in which the decision of whether to change levels and which level to change to was only made after each block of 10 trials, rather than after each trial. At the end of each block, N(C) was calculated on the basis of the 10 trials in the block, and compared to E[N(C)], which was 7.5 (i.e., 75%). This modification was made because pilot testing with the original PEST algorithm yielded unreliable thresholds and, also, to ensure there was enough samples at each stimulus magnitude to enable psychometric functions to be used to fit the data.

11.3.2 Method

11.3.2.1 Participants

Three participants took part in the experiment. All were males recruited from the Human Movements Laboratory and the Psychology Department at the University of Queensland, with ages ranging from 25-32 (mean age 27.3). Two participants were paid $10 for their participation.

11.3.2.2 Procedure

A single PEST staircase was run for each change type and these staircases were interleaved so that the staircase used in any given trial was chosen at random. Each staircase lasted for 200 trials, so the experiment was 1000 trials long.

11.3.3 Results

As for Experiments 5 and 6, c values were calculated to identify any response bias the participants had. These values are given in Table 11.13. All c values were within 1 standard deviation unit of 0 and so all participants' data were retained for further analysis.

Table 11.13: Experiment 7: c Values for Participants from SDT Analysis

	AB	TW	WM
c	-0.04	-0.29	-0.22

Table 11.14: Experiment 7: Thresholds - Psychometric Functions

Feature Dimension	AB	TW	WM
Colour (h)	1.03	1.34	1.07
Orientation (°)	31.44	576.24*	65.24
Size (°)	2.17	1.92	2.27
Spatial Frequency (cpd)	2.65	2.74	2.45
Speed (° s^{-1})	1.10	1.59	1.80

Psychometric function curves were plotted for each observer for each target type, using *psignifit* (see Wichmann & Hill, 2001a, 2001b). Thresholds were taken from the 75% correct point and slopes were taken from this point also. Threshold and slope values for the three included observers are listed in Tables 11.14 and 11.16, respectively. Two example psychometric functions are shown in Figures 11.7 and 11.8.

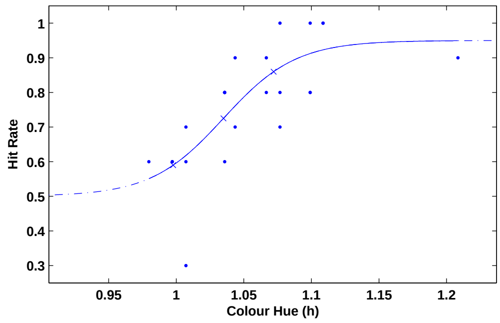


Figure 11.7: A psychometric curve plotted for colour targets for participant AB.

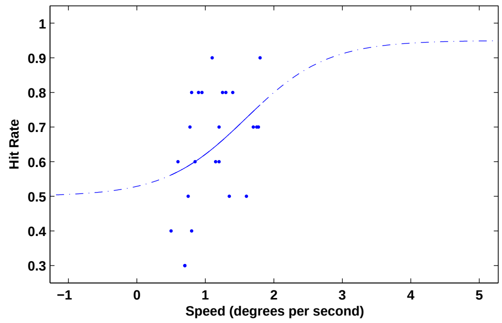


Figure 11.8: A psychometric curve plotted for speed targets for participant TW.

The thresholds in Table 11.14 are relatively consistent across participants, although there is an outlier (marked with an asterisk). This outlier has a much

Table 11.15: Experiment 7: Threshold Confidence Intervals

Feature Dimension	AB	TW	WM
Colour (h)	1.02, 1.04	1.02, 97.70	0.96, 1.12
Orientation (°)	8.77, 39.06	70.99, 977.81	62.03, 258.59
Size (°)	2.13, 2.20	0.94, 2.03	2.20, 2.29
Spatial Frequency (cpd)	-143.75, 2.76	-8.69, 2.77	-62.88, 2.79
Speed (° s^{-1})	1.03, 1.22	1.37, 1.87	NA, NA

Table 11.16: Experiment 7: Slopes

Feature Dimension	AB	TW	WM
Colour (h)	9.33	1.07	1.27
Orientation (°)	0	0	0.02
Size (°)	3.28	-1.69	4.04
Spatial Frequency (cpd)	2.33	-9.20	-0.91
Speed (° s^{-1})	0.78	0.42	145.59

Table 11.17: Experiment 7: Slope Confidence Intervals

Feature Dimension	AB	TW	WM
Colour (h)	6.96, 15.76	0.01, 2.33	0.25, 2.31
Orientation (°)	0, 0.02	0, 0.01	0, 0.04
Size (°)	2.26, 5.17	-3.75, -0.31	1.95, 7.77
Spatial Frequency (cpd)	0, 10.52	-22.28, -0.02	-3.74, 0
Speed (° s^{-1})	0.49, 0.92	0.22, 0.67	NA, NA*

Table 11.18: Experiment 7: Thresholds - Value at Minimum Step Size

Feature Dimension	AB	TW	WM
Colour (h)	0.10	0.14	0.14
Orientation (°)	56	61	60.9
Size (°)	2.13	2.23	2.30
Spatial Frequency (cpd)	1.78	2.82	2.84
Speed (° s^{-1})	2.79	1.88	1.80

Table 11.19: Experiment 7: Thresholds - End of PEST Run

Feature Dimension	AB	TW	WM
Colour (h)	0.04	0.10	0.10
Orientation (°)	29	66	67.5
Size (°)	2.13	2.23	2.30
Spatial Frequency (cpd)	2.77	2.00	1.80
Speed (° s^{-1})	1.35	2.82	2.9

higher variability than the other thresholds and this suggests the data may not have fitted well to a psychometric function. The lack of fitting is also evident in the presence of negative slopes (Table 11.16) and the spread of data points shown in the graphs (Figures 11.7 and 11.8). There are two other methods that are used for calculating thresholds from PEST runs - taking the stimulus value corresponding to the minimum step size and taking the stimulus value at the end of the run (as long as the run is 200 trials or longer) (Leek, 2001). The results of these methods are shown in Tables 11.18 and 11.19.

11.3.4 Discussion

The three different methods (psychometric function, minimum PEST step size, end of PEST run) yielded significantly different thresholds and the psychometric functions demonstrated significant variability in the data. Although the psychometric functions (and hence the thresholds obtained from them) are unreliable, the standard method of extracting thresholds from a PEST run (i.e., the value at the end of the run) may still have yielded accurate thresholds. This can only be checked by making changes at threshold levels and observing the performance of observers. This is done in the next experiment, which uses multiple changes in addition to single changes to compare the effects of different change levels.

11.4 Experiment 8 - Changes to One or Two Objects at Threshold Levels

11.4.1 Introduction

Like Experiment 3, this experiment aimed to examine the detectability of single and multiple changes occurring within and between objects, using a baseline level of change for each change type that produced equivalent levels of detectability across change types. Unlike Experiment 3, this baseline level was unique to each observer, as it was obtained using threshold estimation (in Experiment 7). This should allow a more precise comparison of the different single and multiple change conditions.

11.4.2 Predictions

It was predicted that each observer would produce a similar pattern of performance to that found in Experiment 3. That is, when two features changed performance would be the same regardless of whether the two changes occurred on the same object or across different objects. Also, performance for detecting a change when two changes occurred would be higher than when one change occurred.

11.4.3 Method

11.4.3.1 Participants

Two participants took part in the experiment. Both were males recruited from the Human Movements Laboratory and the Psychology Department at the University of Queensland, with ages ranging from 23-25. One participant was paid $15 for his participation.

11.4.3.2 Conditions

This experiment followed a 2 (target side) x 2 (number of targets) x 15 (target type) design. Stimulus magnitudes for the target features were fixed at the threshold level determined by the last experiment. For each feature other than the target features, half of the elements (randomly chosen) had the lowest value for that featural dimension while the others each had the same value which was randomly chosen from those listed in Table 11.20 at the start of the trial. Although there were only 5 types of features that changed, 10 of the 15 target types involved 2 features changing simultaneously. When there was only one target, these 2 features would change on the same (target) element. When there were two targets, one would change in each target. The 15 target types were generated based on single and double combinations of colour, size, orientation, speed and spatial frequency. A list is given in Table 11.21 along with the results for single object conditions.

Table 11.20: Experiment 8: Stimulus Parameters for Size and Speed.

Feature Dimension	min	max	step size
Orientation (°)	0	84	6
Size (°)	1.80	2.6	0.1
Spatial Frequency (cpd)	1.5	3.1	0.2
Speed (° s^{-1})	0.30	0.86	0.07

11.4.4 Results

Table 11.21 gives proportions correct for single object conditions while Table 11.22 gives them for double object conditions. The results are summarised by means based on the number of objects and features changing, shown in Figures 11.9 and 11.10.

11.4.5 Discussion

One obvious problem with these results is that the proportions correct for single features changing on a single object are, in general, higher than 0.75 (the threshold value) and vary between different features. Given that these proportion correct values were calculated from a high number of trials in each condition, it is likely that the method of threshold establishment is inaccurate. A possible

Table 11.21: Experiment 8: Proportion Correct for Different Feature Combinations for a Single Object

Feature Dimension	AB	TW
Colour (h)	0.875	0.55
Orientation (°)	0.875	0.85
Size (°)	0.7	0.875
Spatial Frequency (cpd)	0.892	0.913
Speed ($^\circ$ s^{-1})	0.917	0.675
Colour (h) and Orientation	0.938	0.725
Colour and Size	0.963	0.875
Colour and Spatial Frequency	0.975	0.925
Colour and Speed	0.938	0.938
Orientation and Size	0.938	0.938
Orientation and Spatial Frequency	0.975	0.975
Orientation and Speed	0.975	0.938
Size and Spatial Frequency	0.938	0.975
Size and Speed	0.95	0.95
Spatial Frequency and Speed	0.942	0.975

Table 11.22: Experiment 8: Proportion Correct for Different Feature Combinations for Two Objects

Feature Dimension	AB	TW
Colour (h)	0.913	0.575
Orientation (°)	0.888	0.8
Size (°)	0.75	0.8
Spatial Frequency (cpd)	0.925	0.925
Speed ($^\circ$ s^{-1})	0.95	0.875
Colour (h) and Orientation	0.875	0.913
Colour and Size	0.975	0.838
Colour and Spatial Frequency	0.938	0.938
Colour and Speed	0.9	0.888
Orientation and Size	0.875	0.875
Orientation and Spatial Frequency	0.963	0.963
Orientation and Speed	0.963	0.913
Size and Spatial Frequency	0.963	0.925
Size and Speed	0.85	0.9
Spatial Frequency and Speed	0.983	0.938

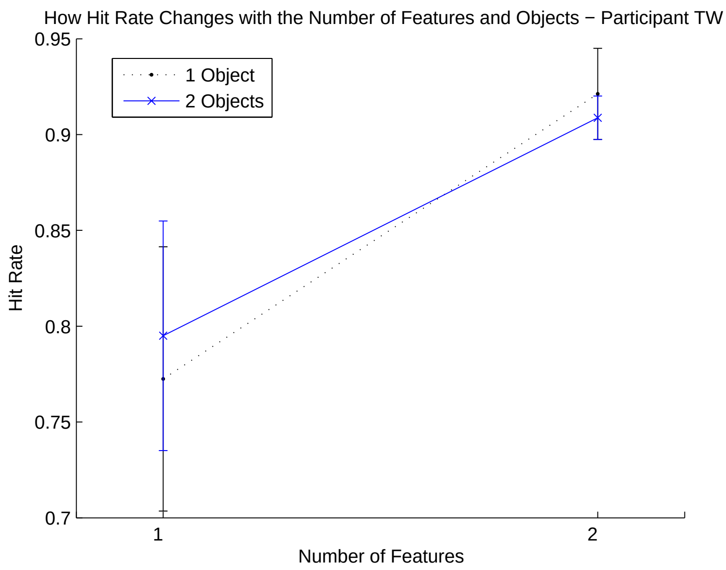


Figure 11.9: How hit rate varies with the number of objects and features changing - TW.

reason for this is that thresholds were established for each feature using a trial sequence involving several interleaved PEST staircases - one for each feature type. Originally, the PEST algorithm was used only for establishing thresholds for a single stimulus type, and so the continuity between trials may be important. In a general sense, perhaps the PEST algorithm is only reliable when applied to simpler tasks than the current one.

The only obvious pattern shown by these results was that changes to two features are more easily detectable than single feature changes, regardless of the number of objects (1 or 2) on which these changes were taking place. However, even this is unclear for the 2 object condition for observer AB. There was a slight difference between the 1 and 2 object conditions when two different features were changing for observer AB, but apart from this, the pattern of data was the same for 1 and 2 object changes. This is different to the pattern of data found in Experiment 3, which also looked at single and multiple changes, between and within objects. More important, perhaps, is the similarity between both experiments - two features changing across two objects produced the same level

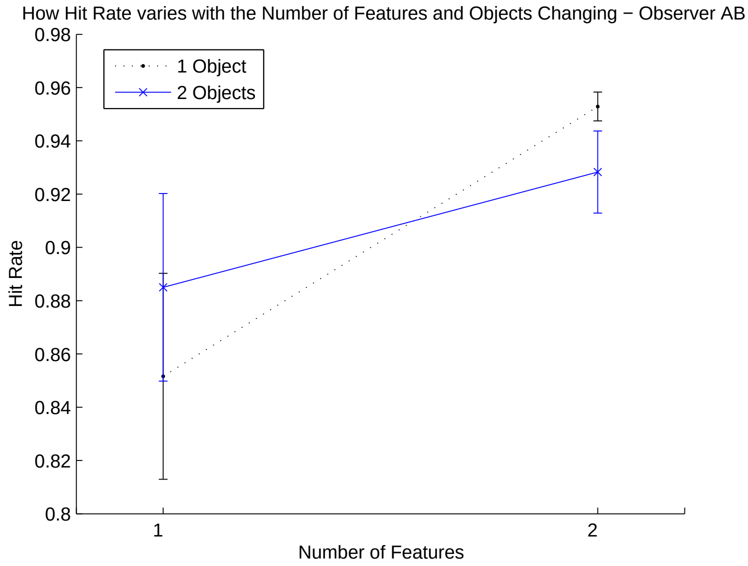


Figure 11.10: How hit rate varies with the number of objects and features changing - AB.

of performance as two features changing within an object. Again, this tends to indicate that featural information as opposed to object information is being selected in these tasks.

Experiment 3 found a difference between single feature changes occurring in one object and single feature changes occurring across two objects, whereas the current experiment did not. However, this difference is only interpretable if the results of the current experiment can be considered valid. This is not necessarily the case, for several reasons. Firstly, the single object, single feature conditions did not yield the hit rates of approximately 0.75 they were expected to. This suggests the PEST method of threshold estimation used in the last experiment was unreliable. This would explain the wide variation between the thresholds estimated using the three different methods (i.e., psychometric function, end of PEST run, minimum step size in PEST run). It would also explain the variability in the slopes and thresholds obtained using the psychometric function method. The reliability of the PEST method is examined in the next experiment.

11.5 Experiment 9 - PEST Runs with the Dummy Observer

11.5.1 Introduction

Because Experiments 7 and 8, taken together, show that it is likely the thresholds obtained using the PEST method are inaccurate, the PEST algorithm itself was tested in the current experiment using an artificial or 'dummy' observer with a known response distribution, so that the accuracy of the algorithm could be thoroughly checked.

11.5.2 The Dummy Observer

There are various ways an adaptive test can be run. It is important that the method chosen accurately generates the stimulus values it is supposed to (i.e., those that correspond to a particular level of performance - for example, 75%). The accuracy of different threshold-obtaining algorithms can be tested by running a dummy observer, with a known response distribution, through a series of trials which use the algorithm to generate stimulus values. In the current study, a dummy observer was created with a range of stimulus values for orientation, size and spatial frequency. For each stimulus value, a corresponding probability value was created. This value represented the probability of the dummy observer making a correct response when that stimulus value is presented. Figures 11.11, 11.13 and 11.12 show the dummy observer's response distributions for size, spatial frequency and orientation, respectively.

On any given trial, the probability of the dummy observer responding correctly to the generated stimulus value is looked up. A random probability value is then generated and, if it is larger than the dummy observer's probability value, an incorrect answer is produced. If it is smaller than or equal to the dummy observer's value, a correct answer is produced. These response distributions are based only on human responses insofar as they have similar 50% and 100% points

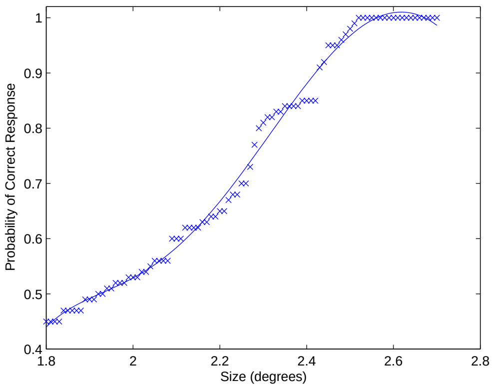


Figure 11.11: Dummy observer response distribution for size.

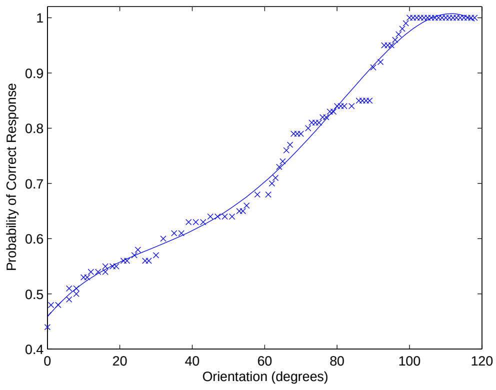


Figure 11.12: Dummy observer response distribution for orientation.

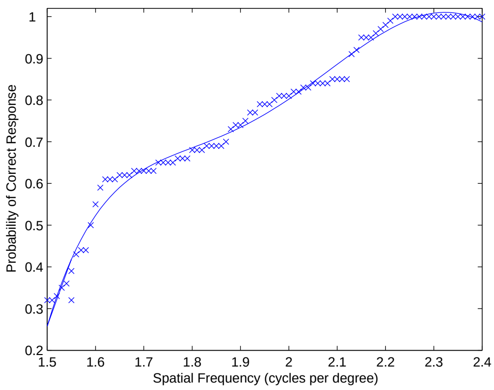


Figure 11.13: Dummy observer response distribution for spatial frequency.

to the experimenter when he ran pilot tests on himself. These points, however vary significantly between observers. The points between the 50% and 100% points in the dummy distributions are not based on any individual's response distribution. Differently shaped distributions were created for each stimulus variable to better test the robustness of the threshold establishment algorithms.

11.5.3 Results of Testing PEST with the Dummy Observer

In Figure 11.14, the mean probability and standard errors are shown for three modified PEST runs of the dummy observer for each of the five change types. It is clear that although most thresholds were close to .75, the variability is high enough to suggest that this is an unreliable method for establishing equivalent levels of detectability across different change types and even, perhaps, within the same change type.

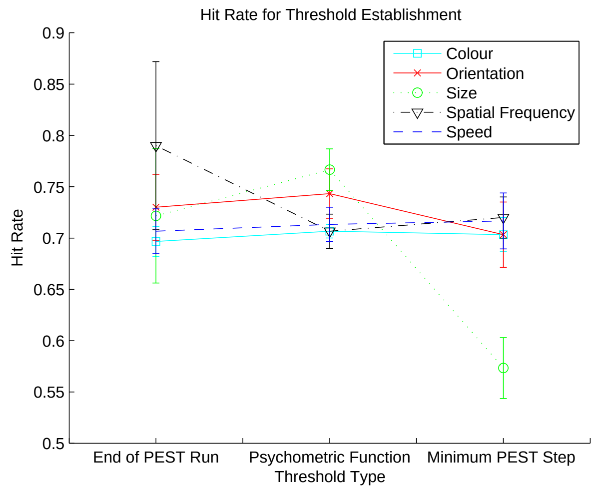


Figure 11.14: The mean probability values and standard errors for the three different threshold establishment methods for the three runs of the dummy observer on each of the five change types.

The variability across change types and the different runs within a change type is greatest for the end of run and minimum step methods. However, even for the psychometric function method, the range of the threshold probability values is more than 0.05, and most of the mean values lie closer to 0.70 than 0.75. It is clear then that the variability of threshold stimulus values from the real observers was not only because of inter-individual variability, but also the unreliability of the algorithm (PEST) used, as a reliable algorithm would generate threshold probability values consistently very close to 0.75.

11.5.4 Conclusion

This experiment shows that the PEST algorithm used for the estimation of thresholds in Experiment 7 was unreliable, at least for the current task. Therefore, Experiment 8 cannot be considered an accurate test of the detection of single and multiple changes. Because of this, the next experiment utilises an alternative method of generating thresholds for the current task.

11.6 Experiment 10 - Thresholds Using Establishment of Range for the Method of Constant Stimuli

11.6.1 Establishment of Range for the Method of Constant Stimuli

Given that the PEST method used in the last Experiment was unreliable and inaccurate in producing thresholds, a new adaptive method was devised. The problem with Experiment 3 was that the ranges of stimulus values that were responded to changed dramatically between observers. Therefore, as the range is established separately for each observer, it should still be possible to then use the method of constant stimuli for establishment of threshold and the psychometric function, once a range has been established. This strategy can be called *Establishment of Range for Method of Constant Stimuli.* An algorithm was devised to attain estimates of the 50% and 100% correct points for each observer. Each of these points was obtained in a separate staircase run. The rules governing the run to determine the 100% correct point were:

- On every incorrect trial, keep the current stimulus value
- On every correct trial, raise the current stimulus value by a fixed amount, which does not change throughout the run

The rules for the run to determine the 50% point were:

- On every incorrect trial, increase stimulus value
- On every correct trial, decrease stimulus value
- On every reversal of direction multiply the current step size (the amount of increase/decrease) by 0.95

The multiplication by 0.95 in the run to establish the 50% point was done to gradually decrease step-size and so 'home-in' on the 50% point. It was found that using smaller multipliers created unreliable estimates.

11.6.2 Running the Dummy Observer on the Establishment of Range Procedure

The dummy observer's results using this algorithm are shown in Table 11.23. Six separate staircases were run consecutively on the dummy observer for each change type. Of these six runs, three were aiming to determine a stimulus value corresponding to a hit rate of .50 and three were aiming to determine a stimulus value corresponding to a hit rate of 1.00.

Table 11.23: Results for the Dummy Observer, with three runs aiming for 50% and 100% hit rates for each change type

Change Type	HR Aim	Stimulus Value	HR	Num. Trials
Orientation (°)	.5	8.75 (2.82)	.512 (0.016)	200 (0)
	1.00	71.33 (18.114)	0.983 (0.017)	98 (2)
Size (°)	.5	1.87 (0.032)	0.473 (0.015)	80 (7.211)
	1.00	2.49 (0.016)	0.973 (0.012)	60 (5.132)
Spatial Frequency (cpd)	0.5	1.62 (0.005)	0.593 (0.012)	93.667 (10.333)
	1.00	2.31 (0)	1 (0)	43 (0)

The mean Hit Rates in Table 11.23 are all close to 0.50 and 1.00 for the 50% and 100% runs, respectively. Therefore, the algorithm is a reliable method of establishing these values. A range of 6 points on each stimulus dimension was established using these values - the first and last point being the points yielded by the 50% and 100% establishment algorithms respectively, and the other points being evenly spaced between them.

11.6.3 Running Real Observers using the Establishment of Range Procedure

The first participant was run using the change types: orientation, size and spatial frequency (as used for the Dummy Observer) and the results of these runs are shown in Table 11.24.

Table 11.24: Results for Observer KM for Establishment of Range

Change Type	HR Aim	Stimulus Value	Num. Trials
Orientation (°)	.5	7.74	200
	1.00	412.00	200
Size (°)	.5	1.88	53
	1.00	3.25	200
Spatial Frequency (cpd)	0.5	1.83	98
	1.00	3.70	200

Because the 100% point for orientation is greater than 180°, it is not meaningful, as angles of rotation greater than 180° produce exactly the same orientation as angles less than 180°. Therefore, it is likely that orientation changes in this experiment are not producing a clear relationship between performance and change magnitude. For this reason, orientation was dropped as a change type and replaced by speed. Therefore, the dummy observer and observer KM were tested with speed changes and Tables 11.25 and 11.26 shows these results.

Table 11.25: Results for the Dummy Observer, with three runs aiming for 50% and 100% hit rates for each change type

Change Type	HR Aim	Stimulus Value	HR	Num. Trials
Speed (° per second)	.5	0.388 (0.03)	.482 (0.013)	47.33 (1.855)
	1.00	1.85 (0)	1.00 (0)	200 (0)

Table 11.26: Results for Participant KM's Establishment of Range Procedure for Speed Changes

Change Type	HR Aim	Stimulus Value	HR	Num. Trials
Speed (° per second)	.5	0.44	.51	51
	1.00	1.85	1.00	200

Given that size, spatial frequency and speed produced a set of reasonable values for 50% and 100% points, these three change types were used in the experiment using real observers. First, the establishment of range procedure was conducted and this was followed by the method of constant stimuli procedure.

11.6.4 Method - Establishment of Range

11.6.4.1 Participants

Three participants took part in the experiment: AB, TW and KM. Their ages ranged from 24 to 26 (mean age 25.33), one was female and two were male. Participation was in exchange for the experimenter participating in the participants' own experiments.

11.6.4.2 Procedure

Participants were run through each of six experimental runs: 50% and 100% establishment runs for each of the three change types: speed, size and spatial frequency. For each change type, the 100% run was conducted first and was followed immediately by the 50% run. The order of change types was counterbalanced across participants.

11.6.4.3 Results

Table 11.27 shows the results of the 50% and 100% runs for the three change types, averaged across all observers. The table also shows the range between the top and bottom values, as well as a step size which was calculated for the next part of the experiment, which used 12 points spaced equally along each stimulus dimension for each observer with the first and last points being the 50% and 100% points, respectively.

Table 11.27: Results for the Establishment of Range Experiment

Change Type	Bottom (50%)	Top (100%)	Range	Step Size
Size (°)	1.91 (0.01)	3.03 (0.17)	1.13 (0.17)	0.09 (0.01)
Spatial Frequency (cpd)	1.73 (0.05)	3.5 (0.12)	1.77 (0.08)	0.15 (0.01)
Speed (°)	0.48 (0.07)	2.5 (0.36)	2.02 (0.31)	0.17 (0.03)

11.6.5 Conclusion

The small standard errors for the bottom (50%) values show that these were relatively consistent across observers, while the larger standard errors for the top (100%) values show that these were less consistent across observers. The lower and upper values obtained are greater than those in Experiment 6, which was also attempting to obtain thresholds for changing targets, but used a series of fixed values to do so. Also, the range obtained in this experiment is much greater, meaning that these values should be more likely to yield an accurate threshold value, as there is a greater space to test and this space has been set based on measuring observers, rather than being pre-set based on assumptions and estimation.

11.7 Experiment 11 - Thresholds Using the Method of Constant Stimuli

11.7.1 Introduction

This experiment used the points generated in Experiment 10 to estimate 75% correct threshold points for each of the participants using the *method of constant stimuli*. The method of constant stimuli involves presenting each of a series of points on a stimulus dimension a fixed number of times, but in a random order, and then constructing a psychometric function and extracting a threshold from the data. The random order means that the stimulus magnitude present on any particular trial is not related to that in the next or preceding trials, and so this method avoids effects of expectation and habituation which could affect the data gained from staircase procedures.

11.7.2 Procedure

For this experiment, a run of trials was done separately for each of the three change types and the order of these runs was counterbalanced across participants. In each run, each of the 13 stimulus points were presented 15 times on each side of the screen. Therefore, there were 390 trials for each run.

11.7.3 Results

Before using human observers, the experiment was run on the dummy observer, using the stimulus points established in the last experiment, and the results are shown in Table 11.28.

Table 11.28: Experiment 11: Results for the 3 runs of the Dummy Observer

Change Type	Threshold (75%)	Actual 75% Point
Size (°)	2.28 (0)	2.28
Spatial Frequency (cpd)	1.91 (0)	1.91
Speed ($^\circ/s$)	0.85 (0.03)	0.88

Table 11.29: Experiment 11: Thresholds

Feature Dimension	AB	TW	KM
Size (°)	2.15	2.06	2.12
Spatial Frequency (cpd)	1.94	1.94	2.04
Speed (° s^{-1})	0.62	0.84	0.33

Table 11.30: Experiment 11: Thresholds Confidence Intervals

Feature Dimension	AB	TW	KM
Size (°)	2.12, 2.17	2.01, 2.11	2.05, 2.18
Spatial Frequency (cpd)	1.89, 1.98	1.88, 2.00	1.92, 2.11
Speed (° s^{-1})	0.58, 0.66	0.76, 0.92	-0.45, 0.57

These results show that the only variation in the threshold estimate is for speed. Similarly, the estimates for size and spatial frequency correspond exactly to the actual 75% correct points, while the speed estimate is slightly lower. Because this method appears reliable and consistent across three runs, only a single run was carried out on each real observer for each change type. Example psychometric functions are shown in Figure 11.15 (for size changes for observer KM) and Figure 11.16 (for spatial frequency changes for observer TW). Thresholds and their confidence intervals are given in Tables 11.29 and 11.30 and the corresponding slopes and their confidence intervals in Tables 11.31 and 11.32.

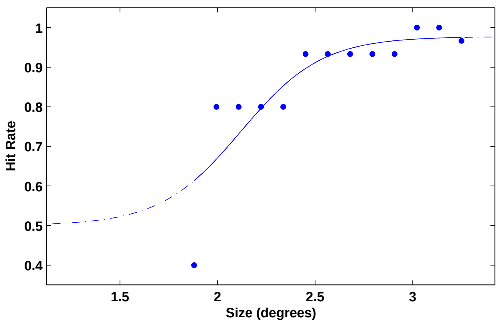


Figure 11.15: A psychometric curve plotted for size targets for participant KM.

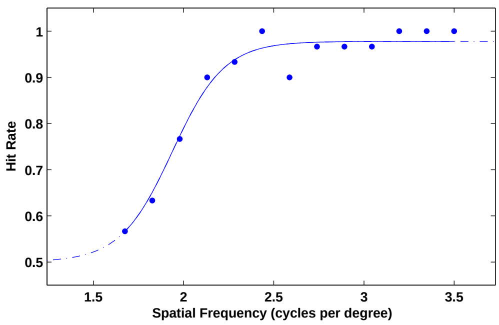


Figure 11.16: A psychometric curve plotted for speed targets for participant TW.

Table 11.31: Experiment 11: Slopes

Feature Dimension	AB	TW	KM
Size (°)	3.66	1.69	1.21
Spatial Frequency (cpd)	2.09	1.73	0.92
Speed (° s^{-1})	2.71	1.09	0.32

Table 11.32: Experiment 11: Slopes Confidence Intervals

Feature Dimension	AB	TW	KM
Size (°)	2.78, 4.79	1.28, 2.13	0.89, 1.66
Spatial Frequency (cpd)	1.54, 2.73	1.18, 2.49	0.68, 1.27
Speed (° s^{-1})	1.88, 3.69	0.80, 1.49	0.20, 0.49

11.7.4 Conclusion

The variability of the thresholds and slopes within observers is fairly small compared to the other experiments that determined thresholds for changing stimuli (experiments 5 and 6). This suggests that the method used was fairly precise in determining thresholds. The threshold values themselves are fairly similar across observers for a given change type, while the slopes show large variability. This could suggest the threshold values are relatively stable for a given change type, but the overall range which elicits a response is not. Although there is clearly much more precision in these threshold measurements than those in Experiment 6, the *accuracy* of the thresholds can only be validated by comparing the detectability of different threshold-level changes, between different observers as well as for different change types for the same observer.

11.8 Experiment 12 - Multiple Changes Using Thresholds Established With the Method of Constant Stimuli

11.8.1 Introduction

Like Experiments 3 and 8, Experiment 12 looked at the detectability of multiple changes occurring simultaneously and compared these against the detectability of single changes. Also like Experiments 3 and 8, this experiment followed up experiments which established some sort of threshold of detectability for a number of different change types. However, the experiments leading up to the current one were significantly more accurate and precise in their methods of threshold establishment than the experiments leading up to Experiments 3 and 8.

11.8.2 Aim

The aim of Experiment 12 was to look at the detectability of multiple changes versus single changes as well as interactions between changes, using two different change types - size and spatial frequency. Originally, speed was to be included as a third change type, but pilot testing indicated performance was at chance level for speed changes.

11.8.3 Method

This experiment used a one-shot paradigm to investigate the detectability of multiple changes. The set size was fixed at six, and there was no cue at the beginning of each trial. Trials proceeded as follows:

$$fixation \rightarrow A \rightarrow B \rightarrow A' \rightarrow response,$$

where fixation was presented for 500 msec, A and A′ for 500 msec each and B for 120 msec. The response screen was presented until the participant responded.

11.8.3.1 Participants

Three participants took part in the experiment. Two were students at the University of Queensland, recruited in exchange for the experimenter's participation in their experiments. The remaining participant was the experimenter himself.

11.8.3.2 Conditions

The experiment followed a 3 (type of change) x 2 (number of objects changing) x 2 (side of change) x 2 (number of features changing) x 5 repetitions. These 96 conditions were repeated 5 times, creating 5 separate blocks of 96 trials each, meaning there was a total of 480 trials for each participant. Trials were arranged in a random order in each block. Changes were fixed at the threshold levels established in Experiment 11. Therefore, each observer responded to different levels of change for the same change type.

11.8.4 Results

Figures 11.17, 11.18 and 11.19 show results for observers AB, KM and TW respectively. For observers AB and TW, changes that occur in two objects result in a greater hit rate than single object changes, while the reverse is true for KM. However, the size of the standard errors in each case suggest that only the differences between 1 and 2 objects for spatial frequency for AB and the differences between 1 and 2 objects for size for TW should be considered.

An obvious problem with these results is that single object changes for size and spatial frequency do not yield a .75 average hit rate as would be expected, given that these changes are occuring at threshold levels (established in Experiment 11). .75 is not even within the bounds of the standard error in most cases where it should be. This suggests that a necessary pre-condition for the results experiment to be interpretable has not been met.

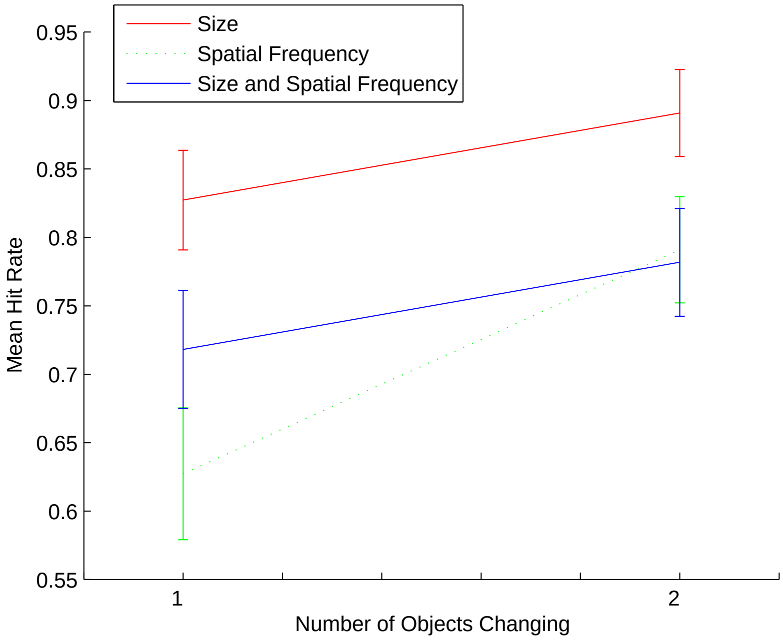


Figure 11.17: Results of Experiment 12 for observer AB.

11.8.5 Discussion

The results of this experiment are inconsistent across observers and do not meet the necessary pre-condition of the experiment, which was that single-feature, single-object changes should yield performance at threshold levels (an average hit rate of 0.75). This is especially surprising given that the method of threshold establishment used for this experiment was fairly rigorous compared to the previous methods used. However, it may simply be the case that more rigour is necessary. More trials could be used in establishing and testing thresholds. Not only would this create a higher probability of 'zeroing in' on the actual threshold of the observer, it would also train them better in the experimental procedure. This would create a higher probability of replication of the same level of performance in the experiment following threshold establishment (i.e., the current experiment). The current experiment was performed several weeks after observers had their thresholds established, due to unavailability of observers. This was not expected to be a problem, however, as observers had a training run before the current experiment.

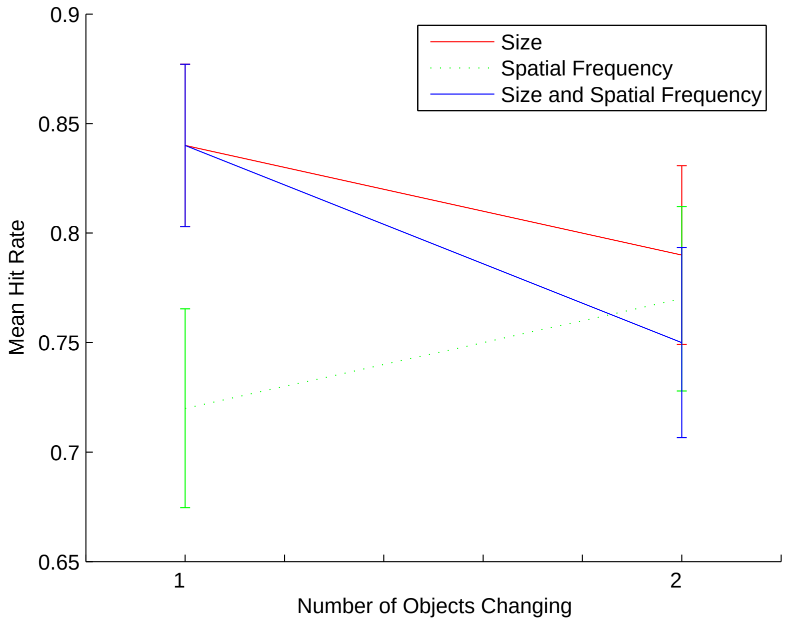


Figure 11.18: Results of Experiment 12 for observer KM.

Experiment 8, like the current experiment, found no overall difference between single and double object changes. However, the current experiment also found no overall difference between single and double feature changes, regardless of whether they took place across 1 or 2 objects. This result was not backed up by Experiment 8 and Experiment 3, which each found differences between 1 and 2 featural changes. This difference may be because the different featural changes in this experiment are equated better than in the other experiments. However, this experiment also found that two features changing in one object was no different (in terms of performance produced) to one object changing one feature. If the salience of the features summated somehow, this would not be expected. It is possible then, that the featural changes interacted in a non-linear fashion or simply did not interact at all. To further examine how featural changes interact, the next experiments use a response paradigm in which the observer responds to each of the changes. In Experiment 13, they are required to select the changing objects and in Experiment 14 they are required to indicate the number of objects that changed.

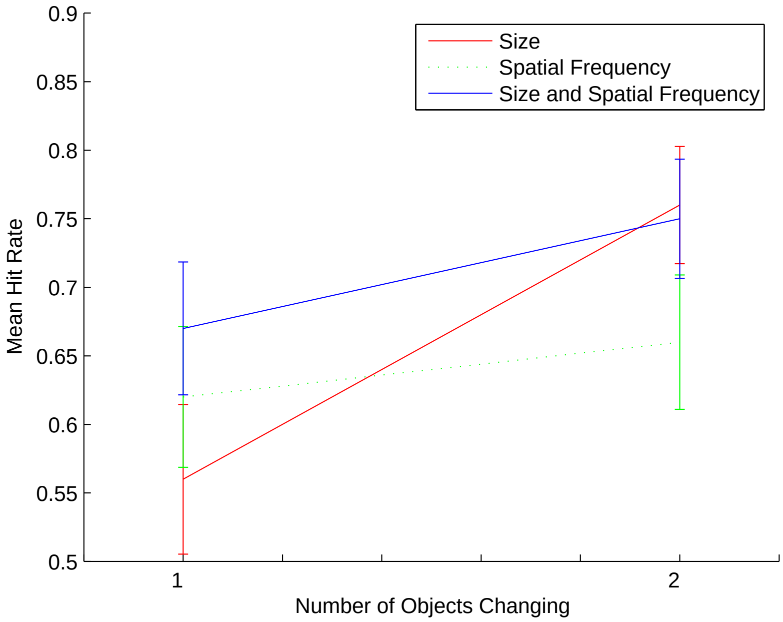


Figure 11.19: Results of Experiment 12 for observer TW.

Chapter 12

Detecting and Identifying Multiple Changes Using Response Paradigms with More Alternatives

12.1 Experiment 13 - Locating Multiple Changes

12.1.1 Introduction

The aim of this experiment was to investigate the detectability of single and multiple changes using a response paradigm that is more precise than the 2AFC methods used previously. The response paradigm used in this experiment requires the participant to select the changing object(s) with the mouse and click on them. Therefore, the participant is required to locate the object(s) undergoing change.

12.1.2 Aim

The aim of Experiment 13 was to look at interactions between different visual features (colour, orientation and spatial frequency), by using a paradigm in which the participant's response gave more information than simply whether they detected a change or not.

12.1.3 Predictions

1. It was expected that the results of the current experiment would support those of Experiment 3 in so far as performance for locating changes would depend on the number of features changing, rather than the number of objects changing

12.1.4 Method

This experiment used a two-shot paradigm to investigate the detectability of multiple changes. The set size was fixed at four, and there was no cue at the beginning of each trial. Trials proceeded as follows:

$$fixation \rightarrow A \rightarrow B \rightarrow A' \rightarrow B \rightarrow A \rightarrow B \rightarrow A' \rightarrow response,$$

where fixation was presented for 1500 msec, A and A′ for 500 msec each and B for 120 msec. The response screen was presented until the participant responded. Figure 12.1 shows an example trial.

12.1.4.1 Participants

Five participants took part in the experiment. All were students at the University of Queensland, recruited through an internet advertisement and each was paid $10 for their participation. Three were female and two were male. Their ages ranged from 18 to 33 and the mean age was 26.

12.1.4.2 Conditions

The experiment followed a 3 (type of change) x 2 (number of objects changing) x 2 (presence of change) x 2 (number of features changing) x 4 (magnitude of change). These 96 conditions were repeated 5 times, creating 5 separate blocks of 96 trials each, meaning there was a total of 480 trials for each participant. Trials were arranged in a random order in each block.

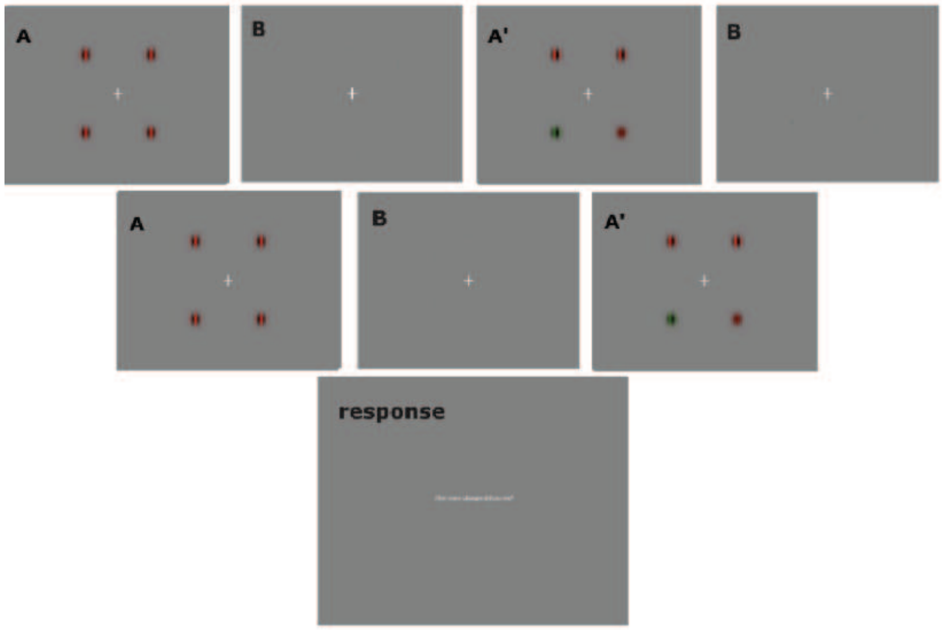


Figure 12.1: An example of a trial in which two objects change, on one feature each. The changing objects are below fixation. The left one changes colour and the right spatial frequency.

12.1.5 Results

Because of the complicated response type, several DVs were created for the experiment. The first - 'totally correct' was 1 if the participant clicked on all the elements that had changed during that trial and 0 if they did not (false positives were ignored). The second - 'partially correct' was 0.5 if the participant clicked on 1 out of 2 changed elements, 1 if they clicked on 1 out of 1 or 2 out of 2 changed elements, and 0 if they clicked on none (false positives were ignored). The third - 'composite score' added the 'partially correct' with a weighting of 0.5 to the 'totally correct' score with a weighting of 1. Results were analysed using three separate ANOVAs. The first used the 'totally correct' DV, the second used the 'partially correct' DV and the third used the 'composite score' DV. Results of these ANOVAs are shown in Tables 12.1, 12.2 and 12.3.

It is clear from the three ANOVAs conducted using different DVs that there is a consistent pattern of results, regardless of the DV used. This pattern shows a main effect of the number of objects changing, an interaction between the

Table 12.1: Experiment 13: ANOVA Totally Correct

Factor	MS	F	p
Number of Features	0.0046	0.743	0.452
Number of Objects	1.2464	20.811	0.002**
Number of Features x Number of Objects	0.1262	11.969	0.008**
Magnitude of Change	1.4048	64.563	0.000**
Magnitude of Change x Number of Features	0.0037	1.532	0.272
Magnitude of Change x Number of Objects	0.3750	17.828	0.000**
Magnitude of Change x Number of Features x Number of Objects	0.0180	3.180	0.026*

Table 12.2: Experiment 13: ANOVA Number Correct

Factor	MS	F	p
Number of Features	0.000	0.005	0.947
Number of Objects	11.556	72.051	0.000**
Number of Features x Number of Objects	0.0834	6.502	0.031*
Magnitude of Change	0.9695	36.992	0.000**
Magnitude of Change x Number of Features	0.0081	0.891	0.482
Magnitude of Change x Number of Objects	0.9695	36.992	0.000**
Magnitude of Change x Number of Features x Number of Objects	0.0049	0.660	0.682

number of features changing and the number of objects changing, a main effect of change magnitude, and an interaction between change magnitude and the number of objects changing. To further analyse these effects, the means of individual conditions are looked at. Figure 12.2 shows the composite score means for the three different number of object conditions for each of the two different feature conditions. This clearly shows the interaction between the number of features changing and the number of objects changing - when one feature is changing, performance is the same regardless of whether that change occurs across one or two objects. However, when two features change, performance is worse when they change across two objects compared to when they both change within one object. This interaction is also reflected in Figure 12.3.

Table 12.3: Experiment 13: ANOVA Composite Score

Factor	MS	F	p
Number of Features	0.008	2.965	0.184
Number of Objects	4.594	104.551	0.000**
Number of Features x Number of Objects	0.050	10.798	0.010*
Magnitude of Change	1.372	65.754	0.000**
Magnitude of Change x Number of Features	0.007	1.585	0.260
Magnitude of Change x Number of Objects	0.354	39.171	0.000**
Magnitude of Change x Number of Features x Number of Objects	0.005	1.493	0.236

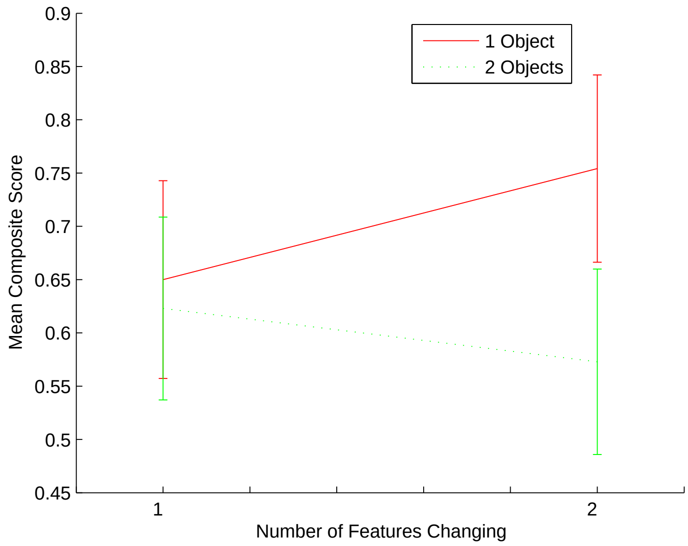


Figure 12.2: Mean composite scores for the number of features changing, for 1 and 2 objects.

12.1.6 Discussion

These results suggest that locating two different changes in two objects is more difficult than locating the same change in two objects, or a change to a single object. Experiment 3 found that the difficulty in detecting single or multiple changes was simply a function of the number of features changing, regardless of how many objects they changed across. However, in the current experiment, it was found that it was more difficult to locate the two objects that both changed, compared to one object changing. This result suggests that object information

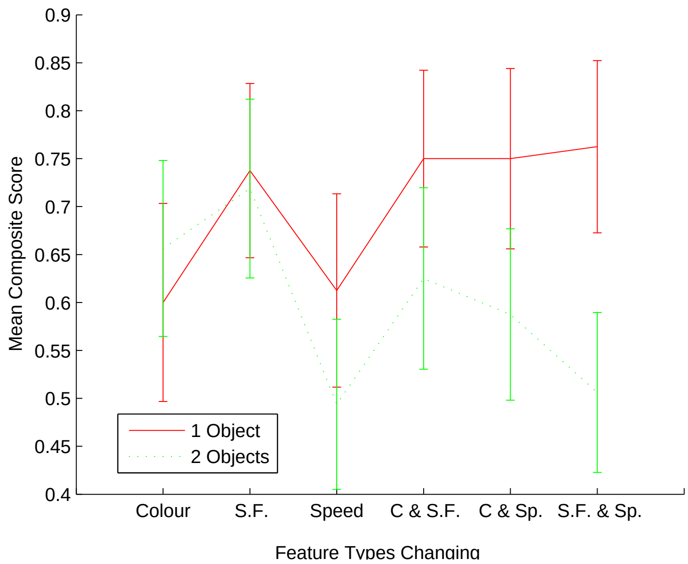


Figure 12.3: Mean composite scores for each of the different feature types that change.

plays a role in the locating of changing objects. While the results of Experiment 3 can be accounted for by a model of feature-based representation, the results of the current experiment cannot. A possible explanation for this is that object information is not required for the detection of change, but is required for the locating of changing items. This would mean that, when the locating response is required, the individual features must somehow become associated with objects in order to facilitate an object-directed response.

Chapter 13

Discussion

One of the underlying tenets of this thesis is that contrary to our subjective experience, our brains do not process the world around us using immediate, exhaustive detail. Instead, evidence from the experiments in this thesis supports the view that our visual system picks out objects and their attributes and represents them in a volatile task-oriented manner. The main focus of the work has been to better understand the content of this representation, specifically the interaction of individual objects and their attributes. To do this, the studies investigated the nature and fidelity of visual short-term memory (vSTM) during scene analysis. In particular the work looked to see if vSTM treats different features differently (i.e., they are subject to different capacity limits) or the same (they represent equivalent, exchangeable tokens). These aims were condensed from a broader theoretical overview discussed in the introduction, which includes the perception of visual scenes, the selection of visual information from visual scenes, and the retention of that information over time. Overall, the thesis aimed to make a contribution to the 'big question' of visual perception, namely: 'How is it that we perceive a stable and coherent world given the limitations, constraints and ambiguity introduced by our sense organs and the external environment?' (see O'Regan, 1992).

The experiments focused on use of the 'change blindness' paradigm, in which participants are required to locate or somehow identify a change that occurs during a given presentation of a scene. This technique has provided a number of important insights into the mechanisms underlying visual scene analysis, but these mechanisms still remain a matter of debate. Using the established techniques of psychophysics and visual search, the thesis attempted to scrutinize the phenomena in a more systematic and rigorous manner than has generally been the case until now (Brown and Orbach (1998) and Rensink (2000c) are notable exceptions).

As part of the effort to simplify and better control the nature of the changes taking place, the experiments used gabor stimuli presented in a visual search paradigm, in which participants searched for a target or several targets amongst distractors. This paradigm yielded several different types of response data. The first, from experiments 1, 2 and 4 concerns the effect of set-size on performance (set-size effects). The second, from experiments 5, 6, 7, 9 and 10, the levels at which performance across different change types are the same - psychometric thresholds. The third, from experiments 3, 8, 12 and 13, the performance for detecting multiple changes versus single changes - multiple target effects. Because the results naturally segment into three general categories, I will first describe them separately, before drawing broader conclusions from the work as a whole.

13.1 Set-size Effects - Experiments 1, 2 and 4

Experiments 1, 2 and 4 used set-size effects as the main dependent measure. Set-size effects look at how performance changes as the number of objects (gabors in this case) which are subject to change increases. Set-size effects can tell the experimenter the cost associated with increasing the number of objects which must be tracked or examined on any trial. Experiments 1, 2 and 4 produced relatively large, consistent effects.

Experiments 1, 2 and 4 used gabors each fixed in a single position, and there were six fixed positions the gabors could take. In Experiments 2 and 4, six gabors were always present, while in Experiment 1, the number of gabors present varied between 1, 2, 4 and 6, but the positions any gabor could take was constrained to the same six positions throughout - and these positions were the same ones used in the other experiments.

Set-size effects in Experiment 1 were fairly similar across change types. In Experiment 2, however, colour changes yielded the smallest set-size effect while the effect for speed was significantly larger and the effect for size was the largest. This colour-speed-size pattern was also shown in Experiment 4, however the magnitudes of each effect were different across the two experiments. Orientation changes exhibited a similar effect size to size changes in Experiment 4, and spatial frequency changes produced an even larger effect. The effect sizes in experiment 4, from smallest to largest, were: colour, speed, size and orientation, spatial frequency.

Set-size effects represent an attempt to measure the limits of the target detection process, that is the range across which change detection is close to perfect (e.g., a set-size of 1) through to a point at which it is close to chance. The detection process itself is presumably affected by a combination of sensory, attentional and memory factors. By its very nature, change detection draws more heavily on memory than a simple search procedure. What the methods in these studies were able to better control was the level of sensory variation. The cued set-size procedure, used in Experiments 2 and 4, made it possible to maintain the number of gabors displayed in each trial, whilst varying the number of patterns within which the subject had to search. Low-level sensory variation was likewise controlled by the 'flicker' (grey blanking period) inherent in the change blindness paradigm which masked the sensory transient created by the changing element. Overall, therefore, one can argue that the sensory input was relatively similar across trials, while the 'attentional input' (governed by cues and instruction) varied. That is not to say that the use of this procedure rules out the influence of sensory factors on the set-size effects, merely that they were more closely controlled than in previous studies and in comparison to the first experiments of the thesis.

The cued set-size paradigm was used by Palmer (1994) (he called it relevant set-size). He made the suggestion, mentioned before, that using this methodology controlled the sensory information present, while varying the attentional information. To justify his claim, he used the results from the conditions using a cued set-size of one, as performance in these conditions should be at ceiling assuming the change is supra-threshold and assuming that observers can selectively attend to the cued elements. In other words, the change should be perfectly detectable in a set-size 1 condition if the amount of sensory information (i.e., the visual presence of 6 gabors) is not relevant to the performance of the task. Palmer's (1994) results backed up his claim - ceiling effects were present for all conditions with a set-size of 1. However, out of the three set-size experiments in this thesis, only Experiment 1 shows an obvious ceiling effect for one-element changes. This suggests that the sensory information did play some role in the detection of changes, as Experiment 1 manipulated sensory information along with attentional information by not using the cued set-size procedure (i.e., the ceiling effect was present only when sensory information was at a minimum). Furthermore, the non-cued elements must have provided a distraction in conditions with a cued set-size of one for the performance to be below ceiling, as the change itself was always supra-threshold. A possible reason for the lack of a ceiling effect in these experiments could be the effectiveness (or lack thereof) of the cue. In Palmer's studies, the cues were presented in the location of the elements, while in the current study, the cue was presented around fixation (as in Brown & Orbach, 1998).

Another reason for the lack of an absolute ceiling for set-sizes of 1 in Experiments 2 and 4 may be because of the increased complexity of the stimuli and task as compared with Palmer's stimuli. Palmer's task was a simple visual search task, with no change. Also, the stimuli were small stationary ovals, presented for a brief period. Furthermore, the distractors were all homogenous and the target differed from the distractors on only one stimulus variable. With the complexity of the stimuli and tasks in the current thesis then, it is perhaps unsurprising that set-sizes of 1 did not generate ceiling effects. However, it is interesting to note that the more controlled conditions of Experiment 4 (compared to Experiment 2) did produce a higher performance level for a set-size of 1. This may be the best that can be done using naive observers in the current paradigm, with the cue being presented at fixation.

Experiment 4 also included an analysis to determine whether the results were best explained by a limited capacity process or an unlimited capacity process. It was found, contrary to expectation, that the small set size effects generated for colour, size and spatial frequency changes were indicative of an unlimited capacity process, while those generated for speed changes were more consistent with a limited capacity process. However, the measurement of these set-size effects was done based on the means of all observers, rather than generating a measure of the effect independently for each observer and then averaging these (as in Palmer, 1998). Also, the fact that the distractors are impairing performance at a cued set size of 1 means that an assumption on which the set-size effect analysis is based is not necessarily being met. The assumption is that the amount of noise present increases in a predictable fashion with set size (as would be the case with a maximally effective cue). Instead, it is clear that noise from the distractors is contaminating detection performance at a set size of 1, and most probably then at all set sizes. Furthermore, it is impossible to tell how much the noise is contributing to the decrement in performance at each set size. To better assess this, it would be necessary to compare set size effects obtained in a cued set size paradigm with those obtained in a display set size paradigm (as in Palmer, 1998).

In summary, the experiments dealing primarily with set-size found a relatively consistent set of results. For change magnitudes that were found to be relatively similar in terms of detectability in pilot testing, set-size effects followed the following ascending pattern: colour, speed, size and orientation, spatial frequency; and this pattern emerged from many observers participating in several experiments. This pattern of results disagrees with Wilken and Ma (2004), who found largest effects for colour changes, smaller effects for spatial frequency changes and smaller effects again for orientation changes. However, in their study, different stimuli were used to study each change type, and each was studied in a different experiment. The comparisons were then made on the basis of model estimates of the noise and variance present in each experiment. In this thesis, it is argued that a comparison between change types using the same stimuli in the same experiment is more likely to be accurate. Regardless of this disagreement, the results of both this study and that of Wilken and Ma (2004) argue in favour of a low-threshold approach to change detection. That is, changes should be viewed in the same manner as detection of more traditional psychophysical stimuli - a continuous process that relies on the collection of evidence for the presence of the stimulus (signal) against a background of noise (see Palmer et al. (2000) for a discussion of high- versus low-threshold models of detection).

13.2 Detection Thresholds - Experiments 5, 6, 7 and 11

In order to more accurately make comparisons between stimulus dimensions, threshold measurements can be taken. This was done in Experiments 5, 6, 7 and 11. Experiment 5 involved detection of a target amongst distractors, where the target was different to the distractors on a single dimension only, and no change occurred. This was done to use a set of basic conditions to establish thresholds for the detection of targets defined by a difference to the distractors along one stimulus dimension only. These basic thresholds were being established to compare against thresholds established in later experiments. The great majority of threshold values in Experiment 5 were lower than those in Experiment 6, when comparing the same stimulus types across experiments. This is to be expected given that simple target detection is an easier task than change detection, all other things being equal (which they were in Experiment 5 and 6).

The thresholds in Experiment 5 were relatively similar across the three different observers. This is probably because the task is a basic search task, involving a very simple and quickly presented stimulus and so the observer-side processes involved in detecting a target are low-level and perceptual in nature, meaning the individual differences will be smaller. However, the thresholds produced in Experiment 6, which involved changing elements, were also similar across observers, with the exception of orientation changes. Given that this experiment involved changing stimuli, this result indicates that reliable and consistent thresholds can be established for change detection and not just conventional target detection. This result agrees with the studies of Brown and Orbach (1998) and Brown et al. (2000), who found reliable and consistent thresholds could be established for luminance increments in simultaneously (AA') as well as successively (A $\rightarrow$ A') presented gabor rings.

Although Experiment 9 found the PEST procedure to be unreliable, casting doubt on the thresholds generated in Experiment 7, it found that there was still a level of precision in the algorithm - most of the thresholds generated were around the same level of performance (70%). With this in mind, it is important to note that the thresholds generated in Experiment 7 are similar to those generated in Experiments 5 and 6 with the exception of speed thresholds - in Experiment 7, they are around double the size of those in Experiments 5 and 6. This could be because of some limitation the PEST procedure has when establishing thresholds for a quantity like speed. Perhaps steps in speed are less perceptible than steps in other quantities, and so the overall threshold is elevated in a stepped procedure.

Regardless of this, the consistency across observers, and the consistency across experiments further validates the procedure of Brown and Orbach (1998) as a means of dealing with change detection in a psychophysical context, and the general idea of using psychophysical methods to measure this phenomenon. It also tells us that the individual differences, although larger than a conventional visual search situation (as in Experiment 5), are still fairly small - at least small enough to suggest that there is an underlying or intrinsic limit to the capacity for storing certain features.

Experiment 11, which used a procedure for establishing thresholds that was built up through data from Experiments 9 and 10, yielded thresholds for size, spatial frequency and speed, using the most reliable threshold estimation method of the entire thesis. Again, these thresholds were consistent across observers and also had values similar to those in Experiments 5, 6 and 7.

The slopes produced in Experiment 6 and 7 contained much greater variation than those in Experiment 5 and Experiment 11. However, there was substantial variation in slopes across observers even in Experiment 11. This indicates that while the stimulus points corresponding to a particular level of performance (75%) were similar in the change-detection threshold experiments, the *overall range* of performance was different for each observer. This could be due to the increased complexity of the task. Having a more complex task would mean that detection performance is less likely to be clustered in a well-defined range, as greater complexity in the task/experimental variables will induce greater noise and variability in the response data.

In summary, it was found that detection thresholds were similar across observers and across experiments reinforcing the notion that it is something intrinsic to the changes, rather than the observers, that is dictating detection performance. Orientation thresholds, however, were not consistent across observers but did show some consistent behaviour across experiments.

13.2.1 Orientation Thresholds

The variation in orientation thresholds in Experiments 6 and 7 is possibly due to invalid thresholds being established, as some of the thresholds are above 90^o, which would be the limit for a symmetrical stimulus such as the gabor. Orientation changes in Experiments 1 and 4 also produced an unusual pattern of results - there was only a small set-size effect and there was no effect of change magnitude. Therefore, it seems that increasing the size of an orientation change does not increase performance for detection of that change. However, Experiment 5 showed that increasing orientation target-distractor difference for a simple one-display visual search task did produce a corresponding increase in performance levels. This agrees with previous research on target-distractor discriminability (Bergen & Julesz, 1983), but the findings for other experiments does not. The absence of a meaningful performance-magnitude relationship for orientation changes may indicate that orientation changes are somehow more complex than the other change types and that, because of this complexity, orientation differences can only be discriminated consistently in a paradigm where only orientation is under consideration, and the difference is obvious (as in Bergen & Julesz, 1983). This could be due to the discrimination of orientation being more difficult and/or the retention of orientation information being more difficult. This was discussed in relation to Experiment 1.

13.2.2 Accuracy and Precision of Threshold Measurement

In order to make valid comparisons across different change types, the thresholds and slopes obtained must be accurate and reliable. The dummy observer, used in Experiments 9, 10 and 11, proved to be a useful way of testing the threshold establishment procedures. Also of use was the confidence intervals for thresholds and psychometric function slopes - the large size of these in Experiment 7 was the reason the dummy observer was created.

It was found that the PEST procedure produced results that did not correspond to the existing 75% correct point for the dummy observer, and had a level of precision that was reasonable, but not good enough for reliable thresholds. The alternative method devised - *Establishment of Range for Method of Constant Stimuli* provided a very reliable (maximum standard error for three runs was 0.03) and precise (hitting the dummy observer's 75% correct point for all three change types) method of establishing thresholds. This is not only shown by the dummy observer statistics when using this algorithm, but also by the results from real observers in Experiment 1 - confidence intervals for slopes and thresholds were much smaller than those in Experiments 6 and 7.

As in Experiment 6, thresholds in Experiment 11 are quite consistent across observers, for a given change type. There is a lot of variation in slopes, however. This also occurred in Experiment 6. Given that these effects are robust, it would appear that the range of performance in this task varies substantially between observers, but the middle level of performance (half-way between chance and perfect - 75%) does not. This may indicate these stimulus conditions can be used to measure robust, absolute thresholds for simple search tasks, as well as for change tasks. Furthermore, it gives more evidence for the idea that the detection of these change types relies more on intrinsic properties of the stimulus than, say, the observer's response bias.

13.3 Multiple Changes

Multiple changes were investigated in Experiments 3, 8, 12, and 13. The levels of change magnitude used in Experiments 3, 8 and 12 were based on some method of threshold establishment in the preceding experiments in each respective case. In Experiment 3, a very approximate method was used, which relied on average values taken from the performance of all participants collectively. In Experiment 8, the thresholds were based on the PEST method of threshold establishment, which was found to be somewhat unreliable in Experiment 9. The levels used in Experiment 12, however, were based on a very accurate and precise method of threshold establishment. Despite the variation in threshold establishment methods, there are some consistent findings across the multiple change experiments.

Experiments 3 and 8 found that a single object change was easier to detect when two features were changing in an object than when only a single feature was changing in an object. Experiment 13 reinforced this finding by showing that a single changing object was easier to locate when two features were changing on it compared to when only one was. From these findings, it can be concluded that two features changing within one object makes it easier to detect a change in that scene than a single feature changing within one object. This is hardly surprising, however. What is more surprising are the findings for changes occurring across two objects.

For changes occurring across two objects, experiments 3, 12 and 13 found that there was no difference in performance for changes that were made to the same feature in each object compared to when different features changed in each object. In Experiment 8, this lack of a difference was found for one observer (AB), but a small difference was found for the other (TW). The means of the two observers, however, showed no difference between two object/one feature and two object/two feature changes. At first glance, this appears to support object-based accounts of information encoding in vSTM. It seems that it is the number of objects changing, and not the specific features changing that are important. However, as mentioned above in relation to single object changes, two features changing on one object are more detectable than a single feature changing on one object. Therefore, in this case, it is the number of features changing that dictates performance. However, all of the experiments, with the exception of Experiment 13, required a response that may have required featural information only - ‘Was there a change or not?’ (Experiment 3), ‘On which side was the change?' (Experiment 12). Experiment 13, however, required participants to locate the changes they saw by clicking on the objects (gabors) that had changed in the preceding trial. Therefore, it is necessary to look more closely at the results from Experiment 13 to get a more complete picture of how the results from this thesis pertain to the object vs. featural encoding dichotomy.

Experiment 13 found that two changes occuring in a single object made that object very easy (relatively speaking) to locate, while two features changing across two objects made those two objects more difficult to locate. This reflects the result summarised above - two featural changes within an object make that object more salient, and two objects each changing the same feature are just as difficult to detect as two objects each changing on a different feature.

The lack of multiple target effects suggests that when two objects changed, the second object was of no help to the observer in terms of detecting the overall change. This could mean that observers are searching only a subset of objects in a serial manner. The question of whether an observer is processing items in a serial or parallel manner is best addressed using the 'redundant targets' paradigm discussed in the Introduction (Section 3.5). This paradigm uses RT as a dependent measure and shows facilitation for multiple targets (i.e., a reduction in RT with added targets) for targets defined by certain attributes. Although the current study used accuracy as the dependent measure, we can still look to see if there is any facilitation (increased accuracy) for this dependent measure. Overall, multiple targets did not appear to facilitate detection performance when the number of features changing is constant (i.e., performance was the same for two objects each changing on a single feature as for one object changing on two features). This would tend to indicate that observers were using a serial search strategy in the experiments of this thesis. However, an alternative explanation is that the time allocated to a display of the elements in the trial is not sufficient to reveal the facilitation effects thought to be reflective of parallel processing (see

Section 3.5). Therefore, RT should be employed as a dependent measure to aid the interpretation these multiple target experiments.

13.4 Capacity of vSTM

Typically, set-size effects have been used to estimate the storage capacity of various parts of the visual system. This tradition of estimating capacity limits is implicitly based on the following assumptions:

- Within the visual system are a number of functional sub-systems that can be delineated according to task specialisation
- These systems have a finite capacity for storing information, and 'hold on' to particular information at a particular time
- The information that these systems are 'holding on' to can be indirectly measured using stimulus-response type experimentation

These assumptions are particularly strong in the research into visual short-term memory. This is because research continually and consistently identifies a number of around 4-5 'objects' that can be 'held on' to for a short period. These 'objects' are so called because they can be a variety of different visual forms, such as numbers, blobs or even real-life objects (usually depicted in photographs or in virtual reality environments). This consistency across different low-level conditions has caused researchers to propose the existence of the higher-level, general purpose store now known as vSTM, which is thought to deal with information at a more abstract level than lower-level systems.

Most of the experiments in the current thesis somehow addressed the question of the capacity of vSTM. Experiments 1-2 and Experiment 4, which measured the effect of set-size on performance, addressed the question in the traditional way, while Experiment 3 and Experiments 8 and 12-13 looked more precisely at the type of information involved in these limits. The way these experiments did this is by varying the number of changes that occurred in any one trial and also the number of objects that changed. It was found consistently that performance was dictated by the number of featural changes taking place, regardless of whether these changes were distributed across objects and, furthermore, irrespective of whether they were changes to the same or to different features. Therefore, the results of this thesis do provide support for the idea that vSTM acts as a more generalised store of abstract information where changing features are apparently represented as inter-changable tokens, rather than being tied to the specific visual, featural or object-based information that created them. However, it is clear that some object-based information is transmitted into vSTM, as the location of those changes is accurately represented, according to the results of Experiment 13.

Given that the current thesis took a signal detection approach to change detection, it would be somewhat misleading or contradictory to offer definite capacity limits for vSTM using, for example, the formulas given by Pashler (H. Pashler, 1988) or Rensink (Rensink, 2000a). Instead, the low-threshold approach suggests that vSTM should be re-conceptualised as something other than a storage space for a definite number of discrete object entities.

13.5 Conclusion

This thesis established a methodology for looking more precisely at various aspects of change detection, in a quantifiable way. Although there were some inconsistencies amongst experimental results, the consistencies found were more frequent and relevant to the questions motivating the thesis than were the inconsistencies. Overall it was found that this psychophysical-search hybrid paradigm is suited to answering questions regarding attention and vSTM and should perhaps be used instead of more traditional paradigms. The findings indicate that vSTM stores information from the visual scene that is both feature- and object-based since performance of search-for-change tasks can not be adequately predicted by the featural or object-based distribution of information in a scene alone.

The thesis also raises the important point that change blindness is not so much of a 'special' phenomenon that it should be considered separately from the attentional processes involved in search fields requiring integration of information across space only (i.e., simple visual search), rather than space and time (i.e., change blindness). Indeed, models that speak of integrated object units were first talked about in relation to static search (see Kahneman et al., 1992). Later theories such as coherence theory (Rensink, 2000a) and object instantiation (FINST) theory (Pylyshyn & Storm, 1988) referred to dynamic search activities as well as the static search situations.

The results of the experiments in this thesis reinforce the notion put forward by coherence theory and FINST theory that static and dynamic search should be dealt with in a unified framework. Indeed, as described in the introduction, no aspect of scene viewing is truly static. The very early work on visual attention dealt more with the overt behaviours associated with scene viewing - eye and head movements - but it has become obvious now that covert behaviours are probably much more important than these and, indeed, that the overt movements are in some sense subservient to these covert behaviours. Obviously, the problem becomes probing and testing the covert behaviours in a way that generates meaningful data and helps inspire sensible theory. A recommendation generated by the integration of results and theory in this thesis, is that both the external environment (e.g., the distribution of featural information across objects) and the internal environment (i.e., as revealed by detection performance) of the observer should be considered, not in isolation from one another, but as a unit that becomes integrated through reciprocal influences in both directions. The current thesis has attempted to take a step towards this goal by addressing shortcomings existing in current approaches to the study of visual attention and memory.

Chapter 14

Appendix 1 - C Code

14.1 Gabor Generation

```
given: int size, frame, en;
 int i, j, z; float f, c, g, c2;
 for (i = 0; i < size; i++) {
    for (j = 0; j < size; j++) {
        f = bg_value*sin(num_cycles*2.0000*3.1415*(double)i/size
        +offset_radians);
        c = f*exp(((size/2-i)*(i-size/2)+(size/2-j)*(j-size/2))/
        (2.625 * size));
        c += bg_value;
        g = 1-(exp(((size/2-i)*(i-size/2)+(size/2-j)*(j-size/2))/
        (2.625 * size)));
        c2 = bg_value*g;
        for (z = 0; z <= 2; z++){
            element[en-1].texture[frame-1][i][j][z] =
            (GLubyte)((element[en-1].colours[z]*c) +
            (1-element[en-1].colours[z])*c2);
            }
    }
 }
```

References

Adelson, E. H., & Bergen, J. R. (1991). Computational models of visual processing. In M. Landy & J. A. Movshon (Eds.), (chap. The plentopic function and the elements of early vision). Massachusettes: MIT Press.

Aginsky, V., & Tarr, M. J. (2000). How are different properties of a scene encoded in visual memory? *Visual Cognition*, *7*, 147–162.

Allport, D. A. (1971). Parallel encoding within and between elementary stimulus dimensions. *Perception and Psychophysics*, *10*, 104–108.

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*(2), 106-111.

Antes, J. R. (1974). The time course of picture viewing. *103*, 62–70.

Averbach, E., & Coriell, A. (1961). Short-term memory in vision. *Bell System Technical Journal*, *40*, 309–328.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, *4*, 829–839.

Beck, J. (1966). Effect of orientation and of shape similarity on perceptual grouping. *Perception Psychophysics*, *1*, 300–302.

Becker, M. W., Pashler, H., & Anstis, S. M. (2000). The role of iconic memory in change-detection tasks. *Perception*, *29*, 273–286.

Bergen, J. R., & Julesz, B. (1983). Rapid discrimination of visual patterns. *IEEE Transactions on Systems, Man and Cybernetics*, *19*, 857–863.

Biederman, I., Mezzanotte, R., & Rabinowitz, J. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*, 143–177.

Broadbent, D. E. (1958). *Perception and communication.* London: Pergammon Press.

Brown, K. C. S., Baker, M., & Orbach, H. S. (2000). Comparison blindness. *Visual Cognition*, *7*, 253–267.

Brown, K. C. S., & Orbach, H. S. (1998). Contrast discrimination, non-uniform patterns and change blindness. *Proceedings of the Royal Society*, *265*, 2159-2166.

Burr, D., Morrone, M., & Ross, J. (1996). Selective suppression of the magnocellular visual pathway during saccades. *Behavioral Brain Research*, *80*, 1–8.

Carpenter, R. H. S. (1988). *Movements of the eyes.* London: Pion.

Carter, E. C., & Carter, R. C. (1981). Color and conspicuousness. *Journal of Optical Society of America*, *71*, 723–729.

Carter, R. C. (1982). Visual search with color. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 127–136.

Ceraso, J. (1985). Unit formation in perception and memory. *Psychology of Learning and Motivation: Advances in Research and Theory*, *19*, 179–210.

Cermack, G. W. (1971). Short-term recognition memory for complex free-form figures. *Psychonomic Science*, *25*, 209–211.

Clifford, C. W. G. (2002). Perceptual adaptation: motion parallels orientation. *Trends in Cognitive Sciences*, *6*, 136–143.

Coltheart, V. (1999). Fleeting memories: Cognition of brief visual stimuli. In V. Coltheart (Ed.), (chap. Introduction: Perceiving and Remembering Brief Visual Stimuli). Cambridge, Massachusetts: MIT Press.

Daniel J. Simons, C. F. C., & Schnur, T. (2002). Evidence for preserved representations in change blindness. *Consciousness and Cognition*, *11*, 78–97.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193-222.

Deutsch, J., & Deutsch, D. (1963). Attention: some theoretical considerations. *Psychological review*, *70*, 80–90.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, *113*, 501–517.

Eriksen, C. W., & Murphy, T. D. (1987). Movement of attention across the visual field: A critical look at the evidence. *Perception & Psychophysics*, *42*, 299–305.

Eriksen, C. W., & Spencer, T. (1969). Rate of information processing in visual perception: Some results and methodological considerations. *Journal of Experimental Psychology Monograph*, *79*(2), 1-16.

Friedman, A. (1979). Framing pictures: the role of knowledge in automized encoding and memory for gist. *Journal of Experimental Psychology: General*, *108*, 316–355.

Graef, P. D., Christiaens, D., & d'Ydewalle, G. (1990). Perceptual effects of scene context on object identification. *Psychological Research*, *52*, 317–329.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Grimes, J. (1996). Perception. vancouver studies in cognitive science. In K. Atkins (Ed.), (chap. On the failure to detect changes in scenes across saccades). New York: Oxford University Press.

Haslam, N., Porter, M., & Rothschild, L. (2001). Visual search: Efficiency continuum or distinct processes? *Psychonomic Bulletin & Review*, *8*(4), 742-746.

Hatfield, G. (1998). Visual attention. In R. D. Wright (Ed.), (chap. Attention in Early Scientific Psychology). New York: Oxford University Press.

Heijden, A. H. C. van der. (1975). Some evidence for a limited cpaacity parallel selfterminating process in simple visual search tasks. *Acta Psychologica*, *39*, 21-41.

Helmholtz, H. v. (1866). *Treatise of physiological optics, vol. iii: trans. 1925 from the third german edition*. New York: Dover.

Henderson, J. M. (1999). High-level scene perception. *Annual Reviews of Psychology*, *50*, 243–271.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*, 498–504.

Henderson, J. M. (2005). Introduction to real-world scene perception. *Visual Cognition*, *12*, 849–851.

Henderson, J. M., & Hollingworth, A. (1998). Eye guidance in reading and scene perception. In G. Underwoord (Ed.), (chap. Eye Movements During Scene Viewing: An Overview). New York: Elsevier.

Holst, E. von, & Mittelstaedt, H. (1950). Das reafferenzprinzip: Wechselwirkungen zwischen zentralnervensystem und peripherie. *Naturwissenschaften*, *37*, 464–476.

Huang, L., & Pashler, H. (2005). Attention capacity and task difficulty in visual search. *Cognition*, *94*, 101-111.

Humphreys, G. W., & Bruce, V. (1989). *Visual cognition: computational, experimental and neuropsychological perspectives*. Hove, England: Lawrence Erlbaum Associates.

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology*, *23*, 420–456.

Irwin, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 307–317.

Irwin, D. E. (1993). Attention and performance xiv: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience. In D. E. Meyer & S. Kornblum (Eds.), (chap. Perceiving an Integrated Visual World). Cambridge, Massachusetts: MIT Press.

Irwin, D. E., & Andrews, R. V. (1996). Attention and performance xvi: Information integration in perception and communication. In T. Inui & J. L. McClelland (Eds.), (pp. 125–155). Cambridge, MA: MIT Press.

Irwin, D. E., Yantis, S., & Jonides, J. (1983). Evidence against visual integration across saccadic eye movements. *Perception & Psychophysics*, *34*, 49–57.

Irwin, D. E., Zacks, J. L., & Brown, J. S. (1990). Visual memory and the perception of a stable visual world. *Perception & Psychophysics*, *47*, 35–46.

James, W. (1890/1950). *The principles of psychology*. New York: Dover.

Johnson, J. S., & Olshausen, B. A. (2003). Timecourse of neural signatures of object recognition. *Journal of Vision*, *3*(7).

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*, 175–219.

Kelley, T. A., Chun, M. M., & Chua, K.-P. (2003). Effects of scene inversion on change detection of targets matched for visual salience. *Journal of Vision*, *3*, 1–5.

Kinchla, R. A., & Wolfe, J. M. (1979). The order of visual processing: 'top-down', 'bottom-up', or 'middle-out'. *Perception & Psychophysics*, *25*, 225–231.

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics*, *63*, 1279–1292.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human*

Perception and Performance, *4*, 565–572.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279-281.

Mack, A., & Rock, I. (1998). *Inattentional blindness.* Cambridge, Massachusetts: MIT Press.

MacKay, D. M. (1973). Visual stability and voluntary eye movements. In R. Jung (Ed.), (pp. 307–331). Berlin: Springer.

Mackworth, N. H., & Morandi, A. J. (1967). The gaze selects informative details within pictures. *Perception & Psychophysics*, *2*, 547–552.

Magnussen, S. (2000). Low-level memory processes in vision. *Trends in Neurosciences*, *23*, 247–251.

Magnussen, S., & Dyrnes, S. (1994). High-fidelity perceptual long-term memory. *Psychological Science*, *5*, 99–102.

Magnussen, S., & Greenlee, M. W. (1992). Retention and disruption of motion information in visual short-term memory. *Journal of Experimental Psychology: Human Learning and Memory*, *18*, 151–158.

Magnussen, S., & Greenlee, M. W. (1997). Competition and sharing of processing resources in visual discrimination. *Journal of Experimental Psychology: Human Learning and Memory*, *23*, 1603–1616.

Magnussen, S., Greenlee, M. W., Asplund, R., & Dyrnes, S. (1991). Stimulus-specific mechanisms of visual short-term memory. *Vision Research*, *31*, 1213–1219.

Mannan, S., Ruddock, K. H., & Wooding, D. S. (1995). Automatic control of saccadic eye movements made in visual inspection of briefly presented 2-d images. *Spatial Vision*, *9*, 363–386.

Marr, D. (1982). *Vision : a computational investigation into the human representation and processing of visual information.* San Francisco: W. H. Freeman.

Maylor, E. A. (1985). Attention and performance vol. 11. In M. I. Posner & O. S. M. Marin (Eds.), (chap. Facilitatory and inhibitory components of orienting in visual space). Hillsdale, NJ: Erlbaum.

McConkie, G. W. (1979). Is visual information integrated across sucessive fixations? *Perception & Psychophysics*, *25*, 221–224.

McConkie, G. W., & Rayner, K. (1976). Identifying the span of the effective stimulus in reading: Literature reviews and theories of reading. In H. Singer & R. B. Ruddell (Eds.), (pp. 137–162). Newark, NJ: International Reading Association.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Moore, C. M., & Egeth, H. (1997). Perception without attention: evidence for grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 339–352.

Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E., & Chabris, C. F. (2001). How not to be seen: The contribution of similarity and selective ignoring to sustained inattentional blindness. *Psychological Science*, *12*, 9–17.

Mullin, P. A., & Egeth, H. E. (1989). Capacity limitations in visual word processing. *15*(3), 111-123.

Murdoch, B. J. (1962). The serial position effect in free recall. *Journal of Experimental Psychology*, *64*, 482–488.

Nagy, A. L., & Sanchez, R. R. (1990). Critical color differences determined with a visual search task. *Journal of Optical Society of AmericanA*, *7*, 1209–1217.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

Neisser, U. (1979). Perception and its development: A tribute to elanor j. gibson. In A. D. Pick (Ed.), (chap. The control of information pickup in selective looking). Hillsdale, NJ: Erlbaum.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology*, *7*, 480–494.

O'Regan, J., Rensink, R., & Clark, J. (1999). Change blindness as a result of 'mudsplashes'. *Nature*, *398*, 34.

O'Regan, J. K. (1992). Solving the real mysteries of visual perception: the world as an outside memory. *Canadian Journal of Psychology*, *46*, 461–488.

Palmer, J. (1990). Attentional limits on the perception and memory of visual information. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 1227–1268.

Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research*, *34*, 1703–1721.

Palmer, J. (1995). Attention in visual search: Distinguishing four causes of a set-size effect. *Current Directions in Psychological Science*, *4*(4), 118-123.

Palmer, J. (1998). Visual attention. In R. D. Wright (Ed.), (chap. Attentional effects in visual search). New York: Oxford University Press.

Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 108–130.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, *40*, 1227–1268.

Pashler, H. (1988). Familiarity and visual change detection. *Perception and Psychophysics*, *44*, 369–378.

Pashler, H. E. (1998). *The psychology of attention.* Cambridge, Massachusetts: MIT Press.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, *16*, 283–290.

Phillips, W. A. (1983). Short-term visual memory. *Philosophical Transactions of the Royal Society of London*, *B302*, 295–309.

Phillips, W. A., & Baddeley, A. D. (1971). Reaction time and short-term visual memory. *Psychonomic Science*, *22*, 73–74.

Phillips, W. A., & Christie, D. F. M. (1977). Components of visual memory. *Quaterly Journal of Experimental Psychology*, *29*, 117–133.

Posner, M. I., & Cohen, Y. (1984). Attention and performance vol. x. In H. Bouma & D. G. Bouwhuis (Eds.), (chap. Components of visual orienting). Hillsdale, NJ: Erlbaum.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the

detection of signals. *Experimental Psychology: General*, *109*, 160–174.
Pylyshyn, Z. (1998). Visual attention. In R. D. Wright (Ed.), (chap. Visual indexes in spatial vision and imagery). New York: Oxford University Press.
Pylyshyn, Z., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*, 1–19.
Quinlan, P. T., & Humphreys, G. W. (1987). Visual search for targets defined by combinations of color, shape and size: An examination of task constraints on feature and conjunction searches. *Perception and Psychophysics*, *41*, 455–472.
Rauschenberger, R. (2003). Attentional capture by auto- and allo-cues. *Psychonomic Bulletin & Review*, *10*, 814–842.
Reisbeck, T. E., & Gegenfurtner, K. R. (1998). Effects of contrast and temporal frequency on orientation discrimination for luminance and isoluminant stimuli. *Vision Research*, *38*, 1105–1117.
Rensink, R. A. (2000a). The dynamic representation of scenes. *Visual Cognition*, *7*, 17-42.
Rensink, R. A. (2000b). The dynamic representation of scenes. *Visual Cognition*, *7*, 17–42.
Rensink, R. A. (2000c). Seeing, sensing and scrutinizing. *Vision Research*, *40*, 1469–1487.
Rensink, R. A. (2000d). Visual search for change: a probe into the nature of attentional processing. *Visual Cognition*, *7*, 345-376.
Rensink, R. A. (2002). Failure to see more than one change at a time. *Journal of Vision*, *2*(7), 245a.
Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*, 368-373.
Riggs, L. A. (1965). Vision and visual perception. In C. H. Graham (Ed.), (pp. 321–349). New York: Wiley.
Ritter, M. (1976). Evidence for visual persistence during saccadic eye movements. *Psychological Research*, *39*, 67–85.
Salamé, P., & Baddeley, A. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior*, *21*, 150–164.
Schriffin, R. M., & Gardner, G. T. (1972). Visual processing capacity and attentional control. *Journal of Experimental Psychology*, *93*, 78-82.
Shaw, M. L. (1980). Attention and performance viii. In R. S. Nickerson (Ed.), (chap. Identifying attention and decision-making components in information processing). Hillsdale, NJ: Erlbaum.
Sherrington, C. S. (1906). *Integrative action of the nervous system*. New Haven, CT: Yale University Press. (reset edition, 1947)
Shore, D. I., & Klein, R. M. (2000). The effects of scene inversion on change detection. *The Journal of General Psychology*, *127*, 27–43.
Shulman, G. L., & Wilson, J. (1987). Spatial frequency and selective attention to local and global information. *Perception*, *16*, 89–101.
Simons, D. J. (2000). Current approaches to change blindness. *Visual Cognition*,

7, 1-15.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception*, *28*, 1059–1074.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present and future. *Trends in Cognitive Sciences*, *9*, 16–20.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, *74*, 1–29.

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *J Comp Physiol Psychol*, *43*, 482–489.

Stevens, J. K., Emerson, R. C., Gerstein, G. L., Kallos, T., Neufield, G. R., Nichols, C. W., et al. (1976). Paralysis of the awake human: Visual perceptions. *Vision Research*, *16*, 93–98.

Tanner, W. P. (1961). Physiological implications of psychophysical data. *Annals of the New York Academy of Sciences*, *89*, 752–765.

Taylor, M., & Creelman, C. D. (1967). Pest: Efficient estimates on probability functions. *The Journal of the Acoustical Society of America*, *41*, 782–787.

Thornton, T., & Gilden, D. L. (2001). Attentional limitations in the sensing of motion direction. *Cognitive Psychology*, *43*, 23-52.

Tipper, S. P., Driver, J., & Weaver, B. (1991). Object-centred inhibition of return of visual attention. *Quarterly Journal of Experimental Psychology*, *43A*, 289–298.

Tipper, S. P., & Weaver, B. (1998). Visual attention. In R. D. Wright (Ed.), (chap. The medium of attention: location-based, object-centered, or scene-based). New York: Oxford University Press.

Tipper, S. P., Weaver, B., Jerreat, L. M., & Burack, A. L. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 478–499.

Townsend, J. T. (1990). Serial vs. parallel processing: Sometimes they look like tweedledum and tweedledee but they can (and should) be distinguished. *Psychological Science*, *1*, 46–54.

Treisman, A. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, *12*, 242–248.

Treisman, A. (1998). Visual attention. In R. D. Wright (Ed.), (chap. The perception of features and objects). New York: Oxford University Press.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.

Treisman, A., Kahneman, D., & Burkell, J. (1983). Perceptual objects and the cost of filtering. *Perception and Psychophysics*, *33*, 527–532.

Turvey, M. T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review*, *80*, 1–52.

Verghese, P. (2001). Visual search and attention: A signal detection theory approach. *Neuron*, *31*, 523–535.

Verghese, P., & Nakayama, K. (1994). Stimulus discriminability and visual search. *Vision Research*, *34*, 2453–2468.

Wallis, G., & Bülthoff, H. (2000). What's scene and not seen: influences of movement and task upon what we see. *Visual Cognition*, *7*, 175-190.

Wallis, G., & Bülthoff, H. (2001). Effects of temporal association on recognition memory. *Proceedings of the National Academy of Sciences*, *98*(8), 4800-4804.

Wheeler, M. E., & Treisman, A. M. (2002). Competition and sharing of processing resources in visual discrimination. *Journal of Experimental Psychology: Human Learning and Memory*, *131*, 1603–1616.

Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. fitting, sampling and goodness-of-fit. *Perception & Psychophysics*, *63*, 1293–1313.

Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: Ii. bootstrap-based confidence intervals and sampling. *Perception & Psychophysics*, *63*, 1314–1329.

Wilken, P. (2001, July). *Capacity limits for the detection and identification of change: Implications for models of visual short-term memory.* (Unpublished PhD Thesis)

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*, 1120–1135.

Wing, A., & Allport, D. A. (1972). Multidimensional encoding of visual form. *Perception and Psychophysics*, *12*, 474–476.

Wolf, W., Hauske, G., & Lupp, U. (1978). How pre-saccadic gratings modify post-saccadic modulation transfer functions. *Vision Research*, *18*, 1173–1179.

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search. *Psychological Science*, *9*, 33–39.

Wolfe, J. M. (1999). Fleeting memories: Cognition of brief visual stimuli. In V. Coltheart (Ed.), (chap. Inattentional Amnesia). Cambridge, Massachusetts: MIT Press.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 419–433.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5*, 1–7.

Yarbus, A. L. (1967). *Eye movements and vision.* New York: Plenum Press.

Yokosawa, K., & Mitsumatsu, H. (2003). Does disruption of a scene impair change detection? *Journal of Vision*, *3*, 41–48.